Class II Human Histone Deacetylases, and Uses Related Thereto

Cross Reference to Related Applications

This application claims the benefit of U.S. Provisional Application No. 60/186,802, filed 3 March 2000, the contents of which are specifically incorporated herein.

Government Funding

Work described herein was supported in part by funding from the National Institute of General Medical Sciences (Grant Number GM38617). The United States Government has certain rights in the invention.

Background of the Invention

The regulation and mechanism of transcription has been of great interest to researchers. In eukaryotic cells, DNA is packaged in the form of nucleosomal arrays. Each nucleosome core consists of 145 base pairs of DNA wound around an octamer of H2A, H2B, H3 and H4 histone proteins. These nucleosome cores are then packaged into higher order structures with additional factors to form chromatin (Luger, K. et al. *Nature,* **1997**, *389*, 251-60).

The incorporation of DNA into chromatin creates a repressive environment that has been implicated in transcriptional silencing. Two cellular processes serve to alter chromatin structure (Workman et al., *Annu. Rev. Biochem.*, **1998**, *67*, 545-579). Chromatin remodeling factors such as SWI/SNF, RSC, NURF, and NRD (reviewed in Varga-Weisz, P.D. and Becker, P.B. *Curr. Opin. Cell Biol.*, **1998**, *10*, 346-353; see also, Tong et al. *Nature*, **1998**, *395*, 917-921; Zhang et al. *Cell,* **1998**, *95*, 279-289; Xue et al., *Mol. Cell*, **1998**, *2*, 851-861) have been shown to increase the accessibility of the DNA, presumably by modifications of the nucleosomal structure. A second cellular mechanism involves alterations of the acetylation state of nucleosomal histones. Hypoacetylated chromatin is often associated with silent genes, while hyperacetylation is correlated with actively transcribed genes. However, this rule is not absolute. Acetylation of K12 on histone H4 is observed in silent heterochromatin regions in *Drosophila* and yeast (reviewed in Grunstein, M. *Nature*, **1997**, *389*, 349-352). Furthermore, there is increasing evidence for regulation of non-histone proteins by acetylation, and this may function in activation as well as repression of transcription (see, Imhof et al., *Curr. Biol.*, **1997**, *7*, 689-692;

Gu, W. and Roeder, R.G. *Cell*, **1997**, *90*, 595-606; Munshi et al. *Mol. Cell*, **1998**, *2*, 457-467). The acetylation state of histones and perhaps non-histone proteins is regulated by a dynamic interaction of histone acetyltransferase (HAT) and histone deacetylase (HDAC) enzymes.

Previously, three human HDACs (Taunton, J. et al. *Science*, **1996**, *272*, 408-411; Yang et al. *J. Biol. Chem.*, **1997**, *272*, 28001-28007; Emiliani et al. *Proc. Natl. Acad. Sci. USA*, **1998**, *95*, 2795-2800; Dagond et al. *Biochem. Biophys. Res. Commun.*, **1998**, *242*, 648-652) and five yeast HDACs (see, Rundlett et al. *Proc. Natl. Acad. Sci USA*, **1996**, *93*, 14503-14508; Carmen et al. *J. Biol. Chem.*, **1996**, *271*, 15837-15844) had been identified and several of these were biochemically characterized. These HDACs, together with the prokaryotic enzymes acetylspermine deacetylase (ASD) and acetoin utilization protein (acuC) comprise a deacetylase superfamily. In yeast, members of this superfamily can be subdivided into two classes based on size and sequence considerations, as well as the observation that Rpd3p and Hda1p function in biochemically distinct complexes. The first class (I) consists of Rpd3p, Hos1p, and Hos2p, while the second class contains Hda1p. Similarly in mammals, HDAC1, HDAC2 and HDAC3 conform to class I criteria, while no human class II HDAC proteins have been identified previously.

Clearly, the identification of alternate classes of genes encoding histone deacetylase proteins would aid in the investigation of functions for these protein products, and thus would be of great benefit in the control of gene transcription and the cell cycle.

Summary of the Invention

In recognition of the desire to understand the cellular function and regulation of histone deacetylases, the present invention, in one aspect, provides heretofore unidentified histone deacetylase genes, and gene products, expressed in mammals. In general, the invention provides a novel class of isolated *HDx* polypeptides, preferably recombinant and/or substantially pure preparations of one or more of the subject *HDx* polypeptides. The invention also provides recombinantly produced *HDx* polypeptides. In other embodiments, the *HDx* polypeptides of the present invention bind to 14-3-3 proteins, such binding resulting in regulation of cellular localization of the *HDx* protein. Additionally, in certain other embodiments, the *HDx* polypeptides of the present invention associated with other *HDx* polypeptides of this novel class and regulate HDAC function by protein pairing

In preferred embodiments, the invention features human class II HDAC nucleic acids and polypeptides. For example, the nucleic acid of SEQ ID NO. 1 encodes HDAC4 and is substantially identical to GenBank Accession No. XM_002252; and the polypeptide of SEQ ID NO. 2 corresponds to HDAC4 and is substantially identical to GenBank Accession No. XP_002252. The nucleic acid of SEQ ID NO. 3 encodes HDAC5 and is substantially identical to GenBank Accession No. XM_008359; and the polypeptide of SEQ ID NO. 4 corresponds to HDAC5 and is substantially identical to GenBank Accession No. XP_008359.2. The nucleic acid of SEQ ID NO. 5 encodes HDAC6 and is substantially identical to GenBank Accession No. NM_006044; and the polypeptide of SEQ ID NO. 6 encodes HDAC6 and is substantially identical to NP_006035.2.

In still other embodiments, the invention features HDAC7 -type human class II HDAC nucleic acids and polypeptides, including both HDAC7A and HDAC7B. HDAC7A is encoded by the nucleic acid of SEQ ID NO. 11 and which is substantially identical to GenBank Accession No. XM_007047. The amino acid sequence of HDAC7A is represented in SEQ ID NO.12 and is substantially identical to GenBank Accession No. XP_007047.1. HDAC7B is encoded by the nucleic acid of SEQ ID NO. 13 and which is substantially identical to GenBank Accession No. XM_004963. The amino acid sequence of HDAC7A is represented in SEQ ID NO.14 and is substantially identical to GenBank Accession No. XP_004963.2.

The *HDx* polypeptides disclosed herein are capable of modulating proliferation, survival and/or differentiation of cells, *inter alia* because of their ability to alter chromatin structure by deacetylating histones. In preferred embodiments the polypeptide has a biological activity including an ability to deacetylate an acetylated histone substrate, preferably a substrate analog of histone H3 and/or H4. In other embodiments the *HDx* polypeptides of the present invention bind to trapoxin or to trichostatin, such binding resulting in the inhibition a deacetylase activity of the *HDx* polypeptide. However, *HDx* polypeptides which specifically antagonize such activities, such as may be provided by dominant negative mutants, are also specifically contemplated.

In addition to acting as a deacetylating enzymes, the *HDx* polypeptides of the present invention are also involved in a novel mechanism for controlling the activity of HDAC proteins. In certain embodiments, polypeptides of the present invention are able to interact with 14-3-3 proteins, whereby cellular localization is regulated. Moreover, in preferred embodiments, the

subject *HDx* proteins have the ability to modulate cell growth by influencing cell cycle progression or to modulate gene transcription.

In one embodiment, the polypeptide is identical with or homologous to an *HDx* protein. Exemplary *HDx* polypeptides include amino acid sequences represented in any one of SEQ ID Nos. 2, 4 or 6. Related members of the *HDx* family are also contemplated, for instance, an *HDx* polypeptide preferably has an amino acid sequence at least 80% homologous to a polypeptide represented by one of more of the polypeptides designated in SEQ ID: Nos: 2, 4, or 6, though polypeptides with higher sequence homologies of, for example, 85%, 90%, 95% or 98% are also contemplated. In one embodiment, the *HDx* polypeptide is encoded by a nucleic acid which hybridizes under stringent conditions with a nucleic acid sequence represented in one or more of SEQ ID Nos: 1, 3 or 5. Homologs of the subject *HDx* proteins also include versions of the protein which are resistant to post-translation modification, as for example, due to mutations which alter modification sites (such as tyrosine, threonine, serine or asparagine residues), or which inactivate an enzymatic activity associated with the protein.

The *HDx* polypeptide can comprise a full length protein, such as represented in SEQ ID Nos. 2, 4 or 6, or it can comprise a fragment corresponding to particular motifs/domains (for example, a ν motif such as shown in SEQ ID Nos. 7, 8, 9 or 10), or to arbitrary sizes, e.g., at least 5, 10, 25, 50, 100, 150 or 200 amino acids in length. In preferred embodiments, the polypeptide, or fragment thereof, specifically deacetylates histones. In other preferred embodiments, the *HDx* polypeptide includes at least one ν motif and in certain embodiments includes two ν motifs.

In certain preferred embodiments, the invention features a purified or recombinatnt *HDx* polypeptide having a molecular weight in the range of 80 kDa to 150 kDa. It will be understood that certain post-translational mofdifications, e.g., phosphorylation, prenylation and the like, can increase the apparent molecular weight of the *HDx* protein relative to the unmodified polypeptide chain.

The subject proteins can also be provided as chimeric molecules, such as in the form of fusion proteins. For instance, the *HDx* protein can be provided as a recombinant fusion protein which includes a second polypeptide portion, e.g., a second polypeptide having an amino acid sequence unrelated (heterologous) to the *HDx* polypeptide, e.g., the second polypeptide portion is glutathione-S-transferase, e.g., the second polypeptide portion is an epitope tage.

In yet another embodiment, the invention features a nucleic acid encoding an *HDx* polypeptide, or polypeptide homologous thereto, which polypeptide has the ability to modulate, e.g., either mimic or antagonize, at least a portion of the activity of a wild-type *HDx* polypeptide. Exemplary *HDx*-encoding nucleic acid sequences are represented by SEQ ID Nos: 1, 3 or 5.

In another embodiment, the nucleic acid of the present invention includes a coding sequence which hybridizes under stringent conditions with one or more of the nucleic acid sequences in SEQ ID. Nos. 1, 3 or 5. The coding sequence of the nucleic acid can comprise a seuqnce which is identical to a coding sequence represented in one of SEQ ID Nos. 1, 3 or 5 , or it can merely be homologous to one or more of these sequences. In preferred embodiments, the nucleic acid encodes a polypeptide which specifically modulates, by acting as either an agonist or antagonist, the enzymatic activity of an *HDx* polypeptide.

Furthermore, in certain preferred embodiments, the subject *HDx* nucleic acid will include a transcriptional regulatory sequence, e.g., at least one of a transcriptional pomoter or transcriptional enhancer sequence, which regulatory sequence is operably linked to the *HDx* gene sequence. Such regulatory sequences can be used to render the *HDx* gene sequence suitable for use as an expression vector. This invention also contemplates the cells transfected with said expression vector whether prokaryotic or eukaryotic and a method for producing *HDx* proteins by employing said expression vectors.

In yet another embodiment, the nucleic acid hybridizes under stringent conditions to a nucleic acid probe corresponding to at least 12 consecutive nucleotides of either sense or antisense sequence of one or more of SEQ ID Nos.: 1, 3 or 5; and more preferably to at least 20. 30, 40 or 50 consecutive (e.g., contiguous) nucleotides; and more preferably to at least 100, 200 or 300 consecutive nucleotides of either sense or antisense sequence of one or more of SEQ ID Nos. 1, 3, or 5.

Yet another aspect of the present invention concerns an immunogen comprising an *HDx* polypeptide in an immunogenic preparation, the immunogen being capable of eliciting an immune response specific for an *HDx* polypeptide; e.g. a humoral response, e.g. an antibody response; e.g. a cellular response. In preferred embodiments, the immunogen comprising an antigenic determinant, e.g. a unique determinant, from a protein represented by one of SEQ ID Nos. 2, 4, 6, 7, 8, 9 or 10.

A still further aspect of the present invention features antibodies and antibody preparations specifically reactive with an epitope of the *HDx* immunogen.

The invention also features transgenic non-human animals, e.g., mice, rats, rabbits, chickens, frogs or pigs, having a transgene, e.g., animals which include (and preferably express) a heterologous form of an *HDx* gene described herein, or which misexpresses an endogenous *HDx* gene, e.g., an animal in which expression of one or more of the subject *HDx* proteins is disrupted. Such a transgenic animal can serve as an animal model for studying cellular and tissue disorders comprising mutated or mis-expressed *HDx* alleles or for use in drug screening.

The invention also provides a probe/primer comprising a substantially purified oligonucleotide, wherein the oligonucleotide comprises a region of nucleotide sequence which hybridizes under stringent conditions to at least 12 consecutive nucleotides of sense or antisense sequence of SEQ ID Nos: 1, 3 or 5, and more preferably to at least 30, 40, 50, 100, 200 or 300 contguous nucleotides of said sequences.

Nucleic acid probes which are specific for each of the *HDx* proteins are contemplated by the present invention, e.g., probes which can discern between nucleic acid encoding a human or bovine HD. In preferred embodiments, the probe/primer further includes a label group attached thereto and able to be detected. The label group can be selected, e.g., from a group consisting of radioisotopes, fluroescent compounds, enzymes, and enzyme co-factors. Probes of the invention can be used as a part of a diagnostic test kit for identifying dysfunctions associated with mis-expression of an *HDx* protein, such as for detecting in a sample of cells isolated from a patient, a level of a nucleic acid encoding a subject *HDx* protein; e.g., measuring an *HDx* mRNA level in a cell, or determining whether a genomic *HDx* gene has been mutated or deleted. These so called "probes/primers" of the invention can also be used as a part of "antisense" therapy which refers to administration or in situ generation of oligonucleotide probes or their derivatives which specifically hybridize (e.g., bind) under cellular conditions, with the cellular mRNA and/or genomic DNA encoding one or more of the subject *HDx* proteins so as to inhibit expression of that protein, e.g., by inhibiting transcription and/or translation. Preferably, the oligonucleotide is at least 12 nucleotides in length, though primers of 25, 40, 50, or 75 nucletotides in length are also contemplated.

In yet another aspect, the invention provides an assay for screening test compounds for inhibitors, or alternativelly, potentiators, or an interaction between an *HDx* protein and an *HDx*

binding protein or nucleic acid sequence. An exemplary method includes the steps of (I) combining an *HDx* polypeptide or fragment thereof, an *HDx* target polypeptide (such as a histone, a 14-3-3 protein, a MEF2 transcription factor, a retinoblastoma associtated protein such as RbAp48, or fragment thereof which interacts with the *HDx* protein), and a test compound, e.g., under conditions wherein, but for the test compound, the *HDx* protein and the target polypeptide are able to interact; and (ii) detecting the formation of a complex which includes the *HDx* protein and the target polypeptide either by directly quantitating the complex, the deacetylase activity of the *HDx* protein, or by measuring inductive effects of the *HDx* protein. A statistically significant change, such as a decrease, in the formation of a complex in the presence of a test compound (relative to what is seen in the absence of a test compound) is indicative of a modulation, e.g., inhibition, of the interaction between the *HDx* protein and its target polypeptide.

In a particularly preferred embodiment, the invention provides combinatorial chemical libraries of compounds for screening for specific inhibitors of class II Histone Deacetylases (HDACs). The combinatorial libraries are used in the assays of the invention in order to identify HDAC inhibitors, including inhibitors specific to all class II histone deacetylases and inhibitors specific to individual class II histone deacetylases. The combinatorial libraries of the invention comprise a plurality of compounds represented by the structures #1, #2, and #3 shown in Figure 14A and the structure shown in Figure 14B. In preferred embodiments, the plurality of compounds are spatially segregated.

Furthermore, the present invention contemplates the use of other homologs of the *HDx* polypeptides or bioactive fragments thereof to generate similar assay formats. In one embodiment, the drug screening assay can be derived with a fungal homolog of an *HDx* protein, such as RPD3, in order to identify agents which selectively inhibit the fungal histone acetylase and not the human enzyme, e.g., for use as antifungal agents.

Yet another aspect of the present invention concerns a method for modulating one or more of growth, differentation, or survival of a mammalian cell by modulating *HDx* bioactivity, e.g., by inhibiting the deacetylase activity of *HDx* proteins, or disrupting certain protein-protein interactions. In general, whether carried out in vivo, in vitro, or in situ, the method comprises treating the cell with an effective amound of an *HDx* therapetuic so as to alter, relative to the cell in the absence of treatment, at least one of (i) rate of growth; (ii) differentiation, or (iii) survival

of the cell. Accordingly, the method can be carried out with *HDx* therapeutics such as peptide and peptidomimetics or other molecules identified in the above-reference drug screens which antagonize the effects of naturally occurring *HDx* protein on said cell. Other *HDx* therapeutics include antisense constructs for inhibiting expression of *HDx* proteins, and dominant negative mutants of *HDx* proteins which competitively inhibit protein-substrate and/or protein-protein interactions upstream and downstream of the wild-type *HDx* protein.

In an exemplary embodiment the subject method is used to treat tumor cells by antagonizing *HDx* activity and blocking cell cycle progression. In one embodiment, the subject method includes the treatment of testicular cells, so as to modulate spermatogenessis. In another emobidmenbt, the subject method is used to modulate osteogenesis, comprising the treatment of osteogenic cells with an *HDx* polypeptide. Likewise, where the treated cell is a chondrogenic cell, the present method is used to modulate chondrogenesis. In still another embodiment, *HDx* polypeptides can be used to modulate the differentiation of progenitor cells, e.g., the method can be used to cause differentiation of a hematopoietic cells, neuronal cells, or other stem/progenitor cell populations, to maintain a cell in a differentiated state, and/or to enhance the survival of a differentiated cell, e.g., to prevent apoptosis, or other forms of cell death.

In addition to such *HDx* therapeutic uses, anti-fungal agents developed with such screening assays as described herein can be used, for example, as preservatives in foodstuff, feed supplement for pomoting weight gain in livestock, or in disinfectant formaulations for treatment of non-living matter, e.g., for decontaminating hospital equipment and rooms. In similar fashion, assays provided herein will permit selection of deacetylase inhibitors which discriminate between the humand and insect deacetylase enzyemes. Accordingly, the present invention expressly contemplates the use and formulations of the deacetylase inhibitors in insecticides, such as for use in management of insects like the fruit fly. Moreover, certain of the inhibitors can be selected on the basis of inhibitory specificity for plant *HDx*-related activities relative to the mammalian enzymes. Thus, the present invention specifically contemplates formulations of deacetylase inhibitors for agricultural applications, such as in the form of a defoliant or the like.

The present method is applicable, for example, to cell culture technique, such as in the culturing of hematopoietic cells and other cells whose survivial or differentiative state is dependent on *HDx* function. Moreover, *HDx* agonists and antagonists can be used for therapetuic interventino, such as to enhance survival and maintenance of cells, as well as to

influence organogenic pathways, such as tissue patterning and other differentiation processes. In an exemplary embodiment, the method is practiced for modulating, in an animal, cell growth, cell differentiation or cell survival, and comprises administering a therapeutically effective amount of an *HDx* polypeptide to alter, relative to the absence of *HDx* treatment, at least one of (I) rate of growth; (ii) differentiation; or (iii) survival of one or more cell-types in the animal.

Another aspect of the present invention provides a method of determining if a subject, e.g., a human patient, is at risk for a disorder characterized by unwanted cell proliferation or aberrant control of differentiation. The method includes detecting, in a tissue of the subject, the presence or absence of a genetic lesion characterized by at least one of (I) a mutation of a gene encoding an *HDx* protein, e.g., represented in one of SEQ ID Nos. 2, 4, 6, 7, 8, 9, or 10 or a homolog thereof; or (ii) the mis-expression of an *HDx* gene. In preferred embodiments, detecting the genetic lesion includes ascertaining the existence of at least one of: a deletion of one or more nucleotides from an *HDx* gene; an addition of one or more nucleotides to the gene, a substitution of one or more nucleotides of the gene, a gross chromosomal rearrangement of the gene; an alteration in the level of a messenger RNA transcript of the gene; the presence of a non-wild type splicing pattern of a messenger RNA transcript of the gene; or a non-wild type level of the protein.

For example, detecting the genetic lesion can include (I) providing a probe/primer including an oligonucleotide containing a region of nucletoide sequence which hybridizes to a sense or antisense sequence of an *HDx* gene, e.g., a nucleic acid represented in one of SEQ ID Nos. 1, 3 or 5 or naturally occurring mutants thereof, or 5' or 3' flanking sequences naturally associated with the *HDx* gene; (ii) exposing the probe/primer to nucleic acid of the tissue; and (iii) detecting, by hybridization of the probe/primer to the nucleic acid, the presence or absence of the genetic lesion; e.g. wherein detecting the lesion comprises utilzing the probe/primer to determine the nucleotide sequence of the *HDx* gene and, optionally, of the flanking nucleic acid sequences. For instance, the probe/primer can be employed in a polymerase chain reaction (PCR) or in a ligation chain reaction (LGR). In alternate embodiments, the level of an *HDx* protein is detected in an immunoassay using an antibody which is specifically immunoreactive with the *HDx* protein.

Description of the Drawing

Figure 1 Predicted amino acid sequences of human class II histone deacetylases. The conserved residues of the catalytic domains are highlighted. (A) HDAC4 (B) HDAC5 (C) HDAC6 predicted amino acid sequences. Note that there are two putative catalytic domains in HDAC6. (D) Alignment of catalytic domains of yeast HDA1p, human HDAC1, 4, and 5 with both catalytic regions of HDAC6. The residues that are conserved in these HDACs as well as in acuC (*B. subtilis*, Accession 348052) and ASD (*M. ramosa*, Accession 3023317) are in bold type, and those residues that are conserved within the Class II human HDAC enzymes are boxed.

Figure 2. Expression analysis of novel human Class II HDAC family members. Multiple human tissue northern blots were probed to determine mRNA expression of HDAC4, HDAC5, and HDAC6. Blots were stripped and reprobed with β-actin cDNA to normalize for total mRNA. The position of molecular size markers is indicated to the left.

Figure 3. Class II HDAC enzymes deacetylate all four core histones *in vitro*. Recombinant FLAG-tagged HDACs were immunoprecipitated from transfected Jurkat cell extracts using α-FLAG antibody (Sigma). Immunopurified enzymes were incubated with radiolabeled core histones as described in Methods. (A) The HDAC activity was measured by scintillation counting of the released $[^3H]$-acetic acid. Where indicated, immunoprecipitates were preincubated with trichostatin A (Wako) prior to addition of histones. Each assay was performed in duplicate and averaged. (B) Substrate specificity of class II HDACs. Deacetylase reactions were separated by 20% SDS/PAGE and stained with Coomassie (top). The gel was treated with EnHance (National Diagnostics), dried and exposed to film (bottom). The identities of the core histones are indicated to the left. RbAp48 was transfected as a negative control.

Figure 4. The catalytic domains of HDAC6 function independently. The histidine residues homologous to H141 of HDAC1 in each of the catalytic domains (H216 and H611) were mutated to alanine by PCR overlap extension. The single and double mutants were FLAG-tagged and expressed in Tag-Jurkat cells. The enzymes were immunoprecipitated using α-FLAG antibodies (Sigma) and expression levels were compared by Western blotting (A). The mutant enzymes were then assayed for histone deacetylase activity as before (B).

Figure 5. Class II HDAC enzymes and HDAC1 are in different complexes *in vivo*. Recombinant FLAG-tagged HDACs were precipitated from transfected Jurkat cell extracts using α-FLAG antibody (Sigma), separated by SDS/PAGE and subjected to Western blot analysis. Blots were probed with (A) α-FLAG antibody (Sigma) to determine expression levels and (B) α-CHD4, α-mSin3A, α-MTA, α-HDAC1, α-HDAC3 and α-Rbp48 antibodies to determine if these proteins co-immunoprecipitated with the Class II HDAC enzymes.

Figure 6. Sequence analysis suggests that HDAC enzymes have diverged into two classes. (A) Alignment of human HDAC enzymes 1 through 6 with yeast Rpd3p, Hos1p, Hos2p, Hos3p, *M. ramosa* ASD, and *B. subtilis* acuC reveals the presence of seven conserved regions, whose consensus sequences differ between the two classes. Amino acids are represented by single letter codes; X represents any amino acid while Φ indicates a hydrophobic residue. NF = not found. (B) A phylogenetic analysis suggests that the HDAC enzymes diverged from a common prokaryotic ancestor to form two classes of HDAC proteins. Proteins from three different phyla were examined. Prokaryotic proteins are preceded by (pro), yeast proteins are preceded by y, while human proteins are capitalized. Note that yHos3p does not correlate well with either HDAC class.

Figures 7A and 7B. Association of HDAC4 and HDAC5 with two isoforms of 14-3-3

A) Recombinant, FLAG-tagged HDAC1 and HDAC4 were transiently expressed in TAg Jurkat cells and immunoprecipitated using α-FLAG agarose (Sigma). Mock-transfected cells were used as a negative control. The immunopurified complexes were separated by SDS/PAGE and the proteins were visualized by silver stain. Two novel bands at 30 and 32 kDa in the HDAC4 immunoprecipitate were identified as 14-3-3 β and ε by peptide microsequencing analysis,

B) The association between HDAC4 and HDAC5 with 14-3-3 was confirmed by Western blot analysis. The recombinant FLAG-tagged proteins were immunoprecipitated with α-FLAG agarose and the purified complexes were separated by SDS/PAGE. The presence of HDAC3, 14-3-3 β and ε was confirmed by probing with specific antibodies (Santa Cruz).

Figures 8A and 8B. Nuclear-cytoplasmic shuttling of HDAC4 and HDAC5 is correlated to 14-3-3 expression levels

A) Recombinant HDAC4-EGFP and HDAC5-EGFP were transiently expressed in U2OS cells and the localization of the protein was observed by fluorescence microscopy.

B) Overexpression of 14-3-3 β causes an increased cytoplasmic localization of HDAC4-EGPF. U2OS cells were transiently transfected with HDAC4-EGFP and either a control plasmid (pcDNA3.1, Invitrogen) or myc-tagged 14-3-3 β. Forty-eight hours post-transfection, the cells were fixed with paraformaldehyde and probed with an α-myc antibody (Upstate Biotechnology) and an α-mouse Ig Texas Red conjugated secondary antibody (Sigma). The DNA was stained with Hoechst dye (Molecular Probes).

Figures 9A and 9B. Phosphorylation-dependent binding of 14-3-3 to HDAC4 and HDAC5

A) Association of HDAC4 and HDAC5 with 14-3-3 and HDAC3 is dependent on the phosphorylation state of the proteins. HDAC4-FLAG and HDAC5-FLAG were transiently expressed in TAg Jurkat cells. Forty-eight hours post-transfection, the cells were treated for 1.5 hours with staurosporine and calyculin A. The immunopurified HDAC4-FLAG and HDAC5-FLAG complexes were subjected to Western blot analysis and tested for HDAC activity, as described in the experimental procedures. The HDAC activity was measured by scintillation counting of the released [^{3}H]-acetic acid.

B) The binding of 14-3-3 to HDAC4 prevents interaction with importin α. Forty-eight hours after transfection with HDAC4-FLAG, TAg Jurkat cells were treated with staurosporin or calyculin A for 1.5 hours. HDAC4-FLAG was immunopurified and subjected to Western blot analysis with α-importin α antibodies (Transducin Laboratories).

Figure 10. Schematic representation of HDLP-TSA complex interactions.

The TSA HDAC inhibitor is shown bound to HPLP (C7-C15 of TSA are labeled). The surrounding HDLP residues are labelled and the corresponding HDAC1 active site residues are indicated below in parentheses. The Zn^{2+} cation in the active site of HDLP is showns as a filled circle. Thatched semi-circles indicate van der Waals contacts between hydrophobic protein resides and TSA, while hydrogen bonds are shown as dashed lines.

Figure 11. Proposed catalytic mechanism for deacetylation of acetylated lysine.
HDLP active site residues and their proposed HDAC counterparts (in parantheses) are labelled.

Figure 12. General structural feature of HDAC inhibitors.
The general structural features of HDAC inhibitors include a Cap group which can be varied to optimize inhibition of a particular HDAC or class of HDACs, an aliphatic chain and a functional group which interacts with the active site of the HDAC.

Figure 13. Alterations in residues contacting TSA in human HDACs.
The residues contacting TSA in HDLP are those which are both shaded and labelled with a consensus residue (i.e. P22 and Y91 residues in the rim of channel, G140, F141, F198, and L265 residues in the channel, and H131, 132, D166, D168, H170, D258, and Y297 residues in the active site) is shown. The surrounding resides that differ between class I and class II HDACs are shaded and labelled with an asterisk (*).

Figure 14. HDAC inhibitor small molecule libraries.
The generalized structures of two representative HDAC inhibitor compound libraries are shown.

Figure 15. Representative positive from HDAC6 screen of printed library.
Cy5HDAC was used to screent he printed 1,3-dioxane library. A representative positive (compound 4-P9) from this screen is shown.

Figure 16. Affect of representative compounds on HDAC1 and HDAC6 activity.

Figure 17. Retesting of resynthesized 11-A15.

Definitions

For convenience, certain terms employed in the specification, examples and appended claims are collected here.

As used herein, the "HDx" polypeptides and nucleic acids of the invention include the histone deacetylace ("HDAC") class II polypeptide and nucleic acid sequences disclosed herein and "HDx" and "HDAC" are used interchangably.

As used herein, the term "nucleic acid" refers to polynucleotides such as deoxyribonucleic acid (DNA), and, where appropriate, ribonucleic acid (RNA). The term should also be understood to include, as equivalents, analogs of either RNA or DNA made from nucleotide analogs, and, as applicable to the embodiment being described, single (sense or antisense) and double-stranded polynucleotides.

As used herein, the term "gene" or "recombinant gene" refers to a nucleic acid comprising an open reading frame encoding one of the novel class of *HDx* polypeptides of the present invention, including both exon and (optionally) intron sequences. A "recombinant gene" refers to nucleic acid encoding an *HDx* polypeptide and comprising *HDx*-encoding exon sequences, though it may optionally include intron sequences which are either derived from a chromosomal *HDx* gene or from an unrelated chromosomal gene. Exemplary recombinant genes encoding the subject *HDx* polypeptide are represented in the appended Sequence Listing. The term "intron" refers to a DNA sequence present in a given *HDx* gene which is not translated into protein and is generally found between exons.

As used herein, the term "transfection" means the introduction of a nucleic acid, e.g., an expression vector, into a recipient cell by nucleic acid- mediated gene transfer. "Transformation", as used herein, refers to a process in which a cell's genotype is changed as a result of the cellular uptake of exogenous DNA or RNA, and, for example, the transformed cell expresses a recombinant form of an *HDx* polypeptide or, where anti-sense expression occurs from the transferred gene, the expression of a naturally-occuring form of the *HDx* protein is disrupted.

As used herein, the term "specifically hybridizes" refers to the ability of the probe/primer of the invention to hybridize to at least 15 consecutive nucleotides of an *HDx* gene, such as an *HDx* sequence designated in one of SEQ ID Nos: 1, 3 or 5 , or a sequence complementary thereto, or naturally occurring mutants thereof, such that it has less than 15%, preferably less than 10%, and more preferably less than 5% background hybridization to a cellular nucleic acid (e.g., mRNA or genomic DNA) encoding a protein other than an *HDx* protein, as defined herein. In preferred embodiments, the oligonucleotide probe specifically detects only one of the subject

HDx paralogs, e.g., does not substantially hybridize to transcripts for other *HDx* homologs in the same species. In general, the term "specifically hybridizes" or "specifically detects" further refers to the ability of a nucleic acid molecule of the invention to hybridize to at least approximately 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 50, 100, 150, 200, 300, 350, or 400 consecutive nucleotides of a vertebrate, preferably a HDAC gene. In certain instances the invention provides nucleic acids which hybridize under stringent conditions to a nucleic acid represented by SEQ ID Nos. 1, 3 or 5 or complement thereof or the nucleic acids. Appropriate stringency conditions which promote DNA hybridization, for example, 6.0 x sodium chloride/sodium citrate (SSC) at about 45° C, followed by a wash of 2.0 x SSC at 50°C, are known to those skilled in the art or can be found in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6 or in Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press (1989). For example, the salt concentration in the wash step can be selected from a low stringency of about 2.0 x SSC at 50°C to a high stringency of about 0.2 x SSC at 50°C. In addition, the temperature in the wash step can be increased from low stringency conditions at room temperature, about 22°C, to high stringency conditions at about 65°C. Both temperature and salt may be varied, or temperature and salt concentration may be held constant while the other variable is changed. In a preferred embodiment, an htrb nucleic acid of the present invention will bind to one of SEQ ID Nos. 1, 3, or 5 or complement thereof under moderately stringent conditions, for example at about 2.0 x SSC and about 40° C. In a particularly preferred embodiment, an HDAC nucleic acid of the present invention will bind to one of SEQ ID Nos. 1, 2, 3, or 4 or complement thereof under high stringency conditions. In another particularly preferred embodiment, a HDAC nucleic acid sequence of the present invention will bind to one of SEQ ID Nos. 1, 3 or 5 which correspond to the HDAC cDNA, preferably ORF nucleic acid sequences, under high stringency conditions.

As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of preferred vector is an episome, i.e., a nucleic acid capalbe of extra-chromosomal replication. Preferred vectors are those capable of autonomous replication and/or expression of nucleic acids to which they are linked. Vectors capable of directing the expression of genes to which they are operatively linked are referred to herein as "expression vectors". In general, expression vectors of utility in

recombinant DNA techniques are often in the form of "plasmids" which refer generally to circular double stranded DNA loops which, in their vector form are not bound to the chromosome. In the present specfiication, "plasmid" and "vector" are used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors which serve equivalent functions and which become known in the art subsequently hereto.

"Transcriptional regulatory sequence" is a generic term used throughout the specification to refer to DNA sequences, such as initiation signals, enhancers, and promoters, which induce or control transcription of protein coding sequences with which they are operable linked. In preferred embodiments, transcription of one of the recombinant *HDx* genes is under the control of a promoter sequence (or other transcriptional regulatory sequence) which controls the expression of the recombinant gene in a cell-type which expression is intended. It will also be understood that the recombinant gene can be under the control of transcriptional regulatory sequences which are the same or which are different from those sequences which control transcription of the naturally-occurring forms of *HDx* genes.

As used herein, the term "tissue-specific promoter" means a DNA sequence that serves as a promoter, i.e., regulates expression of a selected DNA sequence operably linked to the promoter, and which effects expression of the selected DNA sequence in specific cells of a tissue, such as cells of hepatic, pancreatic, neuronal or hematopoietic origin. The term also covers so-called "leaky" promoters, which regulate expression of a selected DNA primarily in one tissue, but can cause at least low level expression in other tissues as well.

As used herein "transgenic animal" is any animal, preferably a non-human mammal, bird or an amphibian, in which one or more of the cells of the animal contain heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of recombinant DNA molecule. This molecule may be integrated within a chromosome, or it may be extrachromosomally replicating DNA. In the typical transgenic animals described herein, the transgene causes cells to express a recombinant form of one of the *HDx* proteins, e.g., either

agonistic or antagonistic forms. However, transgenic animals in which the recombinant *HDx* gene is silent are also contemplated, as for example, the FLP or CRE recombinase dependent constructs described below. Moreover, "transgenic animal" also includes those recombinant animals in which gene disruption of one or more *HDx* genes is caused by human intervention, including both recombination and antisense techniques.

The "non-human animals of the invention include, but are not limited to, vertebrates such as rodents, non-human primates, sheep, dog, cow, chickens, amphibians, reptiles, etc. Preferred non-human animals are selected from the rodent family including rat and mouse, most preferably mouse, though transgenic amphibians, such as members of the Xenopus genus, and transgenic chickens can also provide important tools for understanding and identifying agents which can affect, for example, embryogenesis and tissue formation. The invention also contemplates transgenic insects, including those of the genus Drosophila, such as D. melanogaster. The term "chimeric animal" is used herein to refer to animals in which the recombinant gene is found, or in which the recombinant gene is expressed in some but not all cells of the animal. The term "tissue-specific chimeric animal" indicates that one of the recombinant *HDx* genes is present and/or expresssed or disrupted in some tissues but not others.

As used herein, the term "transgene" means a nucleic acid sequence (encoding, e.g., one of the *HDx* polypeptides, or pending an antisense transcript thereto), which is partly or entirely heterologous, i.e., foreign, to the transgenic animal or cell into which it is introduced, or, is homologous to an endogenous gene of the transgenic animal or cell into which it is introduced, but which is designed to be inserted, or is inserted, into the animal's genome in such a way as to alter the genome of the cell into which it is inserted (e.g., it is inserted at a location which differs from that of the natural gene or its insertion results in a knockout). A transgene can include one or more transcriptional regulatory sequences and any other nucleic acid, such as introns, that may be necessary for optimal expression of a selected nucleic acid.

As is well known, genes for a particular polypeptide may exist in single or multiple copies within the genome of an individual. Such duplicate genes may be identical or may have certain modifications, including nucleotide substitutions, additions or deletions, which all still code for polypeptides having substantially the same activity. The term "DNA sequence encoding an *HDx* polypeptide" may thus refer to one or more genes within a particular individual. Moreover, certain differences in nucleotide sequences may exist between individuals

of the same species, which are called alleles. Such allelic differences may or may not result in differences in amino acid sequence of the encoded polypeptide yet still encode a protein with the same biological activity.

The term "nucleotide sequence complementary to the nucleotide sequence set forth in SEQ ID No. x" refers to the nucleotide sequence of the complementary strand of a nucleic acid strand having SEQ ID No. x. The term "complementary strand" is used herein interchangeably with the term "complement". The complement of a nucleic acid strand can be the complement of a coding strand or the complement of a non-coding strand. When referring to double stranded nucleic acids, the complement of a nucleic acid having SEQ ID No. x refers to the complementary strand of the strand having SEQ ID No. x or to any nucleic acid having the nucleotide sequence of the complementary strand of SEQ ID No. x. When referring to a single stranded nucleic acid having the nucleotide sequence SEQ ID No. x, the complement of this nucleic acid is a nucleic acid having a nucleotide sequence which is complementary to that of SEQ ID No. x. The nucleotide sequences and complementary sequences thereof are always given in the 5' to 3' direction.

The term "percent identical" refers to sequence identity between two amino acid sequences or between two nucleotide sequences. Identity can each be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When an equivalent position in the compared sequences is occupied by the same base or amino acid, then the molecules are identical at that position; when the equivalent site occupied by the same or a similar amino acid residue (e.g., similar in steric and/or electronic nature), then the molecules can be referred to as homologous (similar) at that position. Expression as a percentage of homology, similarity, or identity refers to a function of the number of identical or similar amino acids at positions shared by the compared sequences. Expression as a percentage of homology, similarity, or identity refers to a function of the number of identical or similar amino acids at positions shared by the compared sequences. Various alignment algorithms and/or programs may be used, including FASTA, BLAST, or ENTREZ. FASTA and BLAST are available as a part of the GCG sequence analysis package (University of Wisconsin, Madison, Wis.), and can be used with, e.g., default settings. ENTREZ is available through the National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Md. In one embodiment, the percent identity of two sequences can be determined by the GCG program with

a gap weight of 1, e.g., each amino acid gap is weighted as if it were a single amino acid or nucleotide mismatch between the two sequences.

Other techniques for alignment are described in Methods in Enzymology, vol. 266: Computer Methods for Macromolecular Sequence Analysis (1996), ed. Doolittle, Academic Press, Inc., a division of Harcourt Brace & Co., San Diego, California, USA. Preferably, an alignment program that permits gaps in the sequence is utilized to align the sequences. The Smith-Waterman is one type of algorithm that permits gaps in sequence alignments. See Meth. Mol. Biol. 70: 173-187 (1997). Also, the GAP program using the Needleman and Wunsch alignment method can be utilized to align sequences. An alternative search strategy uses MPSRCH software, which runs on a MASPAR computer. MPSRCH uses a Smith-Waterman algorithm to score sequences on a massively parallel computer. This approach improves ability to pick up distantly related matches, and is especially tolerant of small gaps and nucleotide sequence errors. Nucleic acid-encoded amino acid sequences can be used to search both protein and DNA databases.

Databases with individual sequences are described in Methods in Enzymology, ed. Doolittle, *supra*. Databases include Genbank, EMBL, and DNA Database of Japan (DDBJ).

Preferred nucleic acids have a sequence at least 70%, and more preferably 80% identical and more preferably 90% and even more preferably at least 95% identical to an nucleic acid sequence of a sequence shown in one of SEQ ID Nos. of the invention. Nucleic acids at least 90%, more preferably 95%, and most preferably at least about 98-99% identical with a nucleic sequence represented in one of SEQ ID Nos: 1-4 are of course also within the scope of the invention. In preferred embodiments, the nucleic acid is mammalian. In comparing a new nucleic acid with known sequences, several alignment tools are available. Examples include PileUp, which creates a multiple sequence alignment, and is described in Feng et al., J. Mol. Evol. (1987) *25*:351-360. Another method, GAP, uses the alignment method of Needleman et al., J. Mol. Biol. (1970) *48*:443-453. GAP is best suited for global alignment of sequences. A third method, BestFit, functions by inserting gaps to maximize the number of matches using the local homology algorithm of Smith and Waterman, Adv. Appl. Math. (1981) *2*:482-489.

"Homology" and "identity" each refer to sequence similarity between two polypeptide sequences, with identity being a more strict comparison. Homology and identity can each be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When a position in the compared sequence is occupied by the same amino acid residue, then the polypeptides can be referred to as identical at that position; when the equivalent site is occupied by the same amino acid (e.g., identical) or a similar amino acid (e.g., similar in steric and/or electronic nature), then the molecules can be refered to as homologous at that position. A percentage of homology or identity between sequences is a function of the number of matching or homologous positions shared by the sequences. An "unrelated" or "non-homologous" sequence shares less than 40 percent identity, though preferably less than 25 percent identity, with an *HDx* sequence of the present invention.

As used herein, an "*HDx*-related protein" refers to the *HDx* proteins described herein, and other human homologs of those *HDx* sequences, as well as orthologs and paralogs (homologs) of the *HDx* proteins in other species, ranging from yeast to other mammals, e.g., homologous histone deacetylase. The term "ortholog" refers to genes or proteins which are homologs via speciation, e.g., closely related and assumed to have common descent based on structural and functional considerations. Orthologous proteins function as recognizably the same activity in different species. The term "paralog" refers to genes or proteins which are homologs via gene duplication, e.g., duplicated variants of a gene within a genome. See also, Fritch, WM (1970) Syst Zool 19: 99-113.

"Cells", "host cells", or "recombinant host cells" are terms used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A "chimeric protein" or "fusion protein" is a fusion of a first amino acid sequence encoding one of the subject *HDx* polypeptides with a second amino acid sequence defining a domain (e.g. polypeptide portion) foreign to and not substantially homologous with any domain of one of the *HDx* proteins. A chimeric protein may present a foreign domain which is found (albeit in a different protein) in an organism which also expresses the first proetin, or it may be an "interspecies", "intergenic", etc. fusion of protein structures expressed by different kinds of

organisms. In general, a fusion protein can be respresented by the general formula X-*HDx*-Y, wherein *HDx* represents a portion fo the protein which is derived from one of the *HDx* proteins, and X and Y are, independently, absent or represent amino acid sequences which are not related to one of the *HDx* sequences in an organism.

The term "isolated" as also used herein with respect to nucleic acids, such as DNA or RNA refers to molecules separated from other DNAs, or RNAs, respectively, that are present in the natural source of the macromolecule. For example, an isolated nucleic acid encoding one of the subject *HDx* polypeptides preferably includes no more than 10 kilobases (kb) of nucleic acid sequence which naturally immediately flanks the *HDx* gene in genomic DNA, more preferably no more than 5kb of such naturally occurring flanking sequences, and most preferably less than 1.5 kb of such naturally occurring flanking sequence. The term isolated as used herein also refers to a nucleic acid or peptide that is substantially free of cellular material, viral material, or culture medium when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized. Moreover, an "isolated nucleic acid" is meant to include nucleic acid fragments which are not naturally occurring as fragments and would not be found in the natural state.

<u>Detailed Description of the Invention</u>

The positioning of nucleosomes relative to particular regulatory elements in genomic DNA has emerged as a mechanism for managing the association of sequence specific DNA-binding proteins with promoters, enhancers and other transcriptional regulatory sequences. Two modifications to nucleosomes have been observed to influence the association of DNA-binding proteins with chromatin. Depletion of histones H2A/H2B from thenucleosome facilitates the binding of RNA polymerase II (Baer et al. (1983) Nature 301: 482-488) and TFIIIA (Hayes et al. (1992) PNAS 89:1229-1233). Likewise, acetylation of the core histones apparently destabilizes the nucleosome and is thought to modulate the accessibility of transcription factors to their respective enhancer and promoter elements (Oliva et al. (1990) Nuc. Acid Res. 18: 2739-2747; and Walker et al. (1990) J. Biol. Chem. 265: 5622-5746). In both cases, overall histone-DNA contacts are altered.

In one aspect, the present invention concerns the discovery of new class of histone deacetylase genes in mammals, the gene products of which are referred to herein as "histone

deacetylases" or *HDx*'s. Experimental evidence indicates a functional role for the *HDx* gene products as catalysts of the deacetylation of histones in mammalian cells, and accordingly play a role in determining tissue fate and maintenance. In addition, however, other experimental evidence indicates a role that is biochemically distinct from other classes of HDAC (class I), whereby regulation of cellular localization is involved in the control of the transcriptional activity of HDAC proteins.

The new class of histone deacetylaste genes encode at least three different sub-families, e.g., paralogs, and have been identified from the cells of various mammals. The *HDx* gene products described herein are referred to as HD4, HD5 and HD6, and are represented in SEQ ID No.'s 1-10.

As described below, one aspect of the invention pertains to isolated nucleic acids comprising nucleotide sequences encoding *HDx* polypeptides, and/or equivalents of such nucleic acids. The term nucleic acid as used herein is intended to include fragments as equivalents. The term equivalent is understood to include nucleotide sequences encoding functionally equivalent *HDx* polypeptides or functionally equivalent peptides having an acitivity of an *HDx* protein such as described herein. Equivalent nucleotide sequenes will include sequences that differ by one or more nucleotide substitutions, additions, or deletions, such as allelic variants; and will, therefore, include sequences that differ from the nucleotide sequence of the *HDx* cDNA sequences shown in any of SEQ ID Nos: 1, 3 or 5, due to the degeneracy of the genetic code. Equivalents will also include nucleotide sequences that hybridize under stringent conditions (i.e., equivalent to about 20-27°C below the melting temperature (TM) of the DNA duplex formed in about 1M salt) to the nucleotide sequences represented in one or more of SEQ ID Nos: 1, 3, or 5.

Moreover, it will be generally appreciated that, under certain circumstances, it may be advantageous to provide homologs of one of the subject *HDx* polypeptides which function in a limited capacity as one of either an *HDx* agonist (mimetic) or an *HDx* antagonist, in order to promote or inhibit only a subset of the biological activities of the naturally occurring form of the protein. Thus, specific biological effects can be elicited by treatment with a homolog of limited function, and with fewer side effects relative to treatment with agonists or antagonists which are directed to all of the biological activities of naturally occurring forms of *HDx* proteins.

Homologs of the subject *HDx* proteins can be generated by mutagenesis, such as by discrete point mutation(s), or by truncation. For instance, mutation can give rise to homologs

which retain substantially the same, or merely a subset, of the biological activity of the *HDx* polypeptide from which it was derived. Alternatively, antagonistic forms of the protein can be generated which are able to inhibit the function of the naturally occurring formm of the protein, as for example competing with wild-type *HDx* in the binding of a 14-3-3 protein, a MEF2 transcription factor, RbAp48 or a histone. In addition, agonistic forms of the protein may be generated which are constitutively active, or have an altered K_{cat} or K_m for deacetylation reactions. Thus, the *HDx* protein and homologs thereof provided by the subject invention may be either positive or negative regulators of transcription and/or activation.

In general, polypeptides referred to herein as having an activity of an *HDx* protein (e.g., are "bioactive") are defined as polypeptides which include an amino acid sequence corresponding (e.g., identical or homologous) to all or a portion of the amino acid sequences of an *HDx* protein shown in any one or more of SEQ ID Nos. 2, 4, 6, 7, 8, 9 or 10, which mimic or antagonize all or a portion of the biological/biochemical activities of a naturally occurring *HDx* protein. Examples of such biological activity include the ability to modulate proliferation of cells. For example, inhibiting histone deacetylation causes cells to arrest in G1 and G2 phases of the cell cycle. The biochemical activity associated with *HDx* proteins of the present invention can also be characterized in terms of binding to and (optionally) catalyzing the deacetylation of an acetylated histone.

Other biological activities of the subject new class of *HDx* proteins are described herein or will be reasonably apparent to those skilled in the art. According to the present invention, a polypeptide has biological activity if it is a specific agonist or antagonist of a naturally occurring form of an *HDx* protein.

A. Exemplary Nucleic Acids

Analysis of the new class of *HDx* sequences indicated similarlity with the original class of *HDx* (HD1) and x conserved residues are found. This, along with other experimental data suggests a role in the new class of human *HDx* genes in the deacetylation of histones in mammalian cells. In addition, however, the new class of *HDx* proteins each contains a novel and conserved catalytic region represented by the consensus sequence:

HHAXXXXXXGXCXFNXVAXXAXXXQXXXXXXXXXXLIVDWDXHHGXGTQXXFXXD PSVLYXSXHRYXXGXFXPXX (SEQ ID No.: 7)

With reference to HD4, this catalytic region corresponds to amino acid residues 802-874 of SEQ ID No. 2; with reference to HD5, this catalytic region corresponds to amino acid residues 832-905 of SEQ ID No. 4; and with respect to HD6, this catalytic region corresponds to two catalytic regions corresponding to amino acid residues 215-287 and amino acid residues 610-683 of SEQ ID No. 6.

In another aspect, the invention provides a method of inhibiting a class II *HDx*. The method comprises contacting the *HDx* with a compound capable of initiating *HDx* activity, under conditions such that *HDx* activity is inhibited.

In another aspect, the invention provides a method of purifying an *HDx*. The method includes contacting a reaction mixture comprising an *HDx* with an affinity matrix capable of selectively binding to an *HDx*, and separating at least one other component of the reaction mixture from the *HDx*.

Another aspect of the present invention relates to a method of inducing and/or maintaining a differentiated state, enhancing survival, and/or inhibiting (or alternatively potentiating) proliferation of a cell, by contacting the cells with an agent which modulates *HDx*-dependent transcription. For instance, it is contemplated by the invention that, in light of the present finding of an apparently broad involvement of *HDx* proteins in the control of chromating structure and, thus, transcription and replication, the subject method could be used to generate and/or maintain an array of different tissue both in vitro and in vivo. An "*HDx* therapeutic", whether inhibitory or potentiating with respect to modulatig histone deacetylation, can be, as appropriate, any of the preparations described above, including isolated polypeptides, gene therapy constructs, antisense molecules, peptidomimetics or agents identified in the drug assays provided herein.

In a further embodiment of the invention, the subject *HDx* therapeutics will be useful in increasing the amount of protein produced by a cell or recombinant cell. The cell may include any primary cell isolated from any animal, cultured cells, immortalized cells, and established cell lines. The animal cells used in the present invention include cells which intrinsically have an ability to produce a desired protein; cells which are induced to have an ability to produce a

desired protein, for example, by stimulation with a cytokine such as an interferon, an interleukin; genetically engineered cells into which a gene for a desired protein is introduced. The protein produced by the process could include any peptides or proteins, including peptide hormone or proteinaceous hormones such as any useful hormone, cytokine, interleukin, or protein which it may be desirable to have in purified form and./or in large quantity.

Another aspect of the invention features transgenic non-human animals which express a heterologous *HDx* gene of the present inventino, or which have had one or more genomic *HDx* genes disrupted in at least one of the tissue or cell-types of the animal. Accordingly, the invention features an animal model for developmental diseases, which animal has one or more *HDx* allele which is mis-expressed. For example, a mouse can be bred which has one or more *HDx* alleles deleted or otherwise rendered inactive. Such a mouse model can then be used to study disorders arising from mis-expressed *HDx* genes, as well as for evaluating potential therapies for similar disorders.

Another aspect of the present invention concerns transgenic animals which are comprised of cells (of that animal) which contain a transgene of the present invention and which preferably (though optionally) express an exogenous *HDx* protein in one or more cells in the animal. An *HDx* transgene can encode the wild-type form of the protein, or can encode homologs thereof, including both agonists and antagonists, as well as antisense constructs. In preferred embodimets, the expression of the transgene is restricted to specific subsets of cells, tissues or developmental stages utilizing, for example, cis-acting sequences that control expression in the desired pattern. In the present inventino, such mosaic expression of an *HDx* protein can be essential for many forms of lineage analysis and can additionally provide a means to assess the effects of, for example, lack of *HDx* expression which might grossly alter development in small patches of tissue within an otherwise normal embryo. Toward this end, tissue-specific regulatory sequences and conditional regulatory sequences can be used to control expression of the transgene in certain spatial patterns. Moreover, temporal patterns of expression can be provided by, for example, conditional recombination systems of prokaryotic transcriptional regulatory sequences.

Genetic techniques which allow for the expression of transgenes can be regulated via site-specific genetic manipulation in vivo are know to those skilled in the art. For instance, genetic systems are available which allow for the regulated expression of a recombinase that

catalyzes the genetic recombination a target sequence. As used herein, the phrase "target sequence" refers to a nucleotide sequence that is geneticaly recombined by a recombinase. The target sequence is flanked by recombinase recognition seuqences and is generally either excised or inverted in cells expressing recombinase activity. Recombinase catalyzed recombination events can be designed such that recombination of the target sequence results in either the activation or repression of expression of one of the subject *HDx* proteins. For example, exicsion of a target sequence which interferes with the expression of a recombinant *HDx* gene, such as one which encodes an antagonistic homolog or an antisense transcript, can be designed to activate expression of that gene. This interference with expression of the protein can result from a variety of mechanisms, such as spatial separation of the *HDx* gene from the promoter elemennt or an internal stop codon. Moreover, the transgene can be made wherein the coding sequence of the gene is flanked by recombinase recognition sequences and is initially transfected into cells in a 3' to 5' orientation with respect to the promoter element. In such an instance, inversion of the target sequence will reorient the subject gene by placing the 5' end of the coding sequence in an orientation with respect to the promoter element which allows for promoter driven transcriptional activation.

Similar conditional transgenes can be provided using prokaryotic promoter sequences which require prokayrotic proteins to be simulataneously expressed in order to facilitate expression fo the *HDx* transgene. Exemplary promoters and the corresponding transactivating prokaryotic proteins are given in U.S. Patent No. 4,833,080.

Moreover, expression of the conditional transgenes can be induced by gene therapy-like methods wherein a gene encoding the trans-activating protein, e.g., a recombinase or a prokaryotic protein, is delivered to the tissue and caused to be expressed, such as in a cell-type specific manner. By this method, an *HDx* transgene could remain silent into adulthood until "turned on" by the introduction of the trans-activator.

In an exemplary embodiment, the "transgenic non-human animals" of the invention are produced by introducing transgenes into the germline of the non-human animal. Embryic target cells at various developmental stages can be used to introduce transgenes. Different methods are used depending on the stage of development of the embryonic target cell. The zygote is the best target for micro-injection. In the mouse, the male pronucleus reaches the size of approximately 20 micrometers in diameter which allows reproducible injection of 1-2pl of DNA solution. The

use of zygotes as a target for gene transfer has a major advantage in that in most cases the injected DNA will be incorporated into the host gene before the first cleavage (Brinster et al. (1985) PNAS 82:4438-4442). As a consequence, all cells of the transgenic non-human animal will carry the incorporated transgene. This will in general also be reflected in the efficient transmission of the transgene to offspring of the founder since 50% of the germ cells will harbor the transgene. Microinjection of zygotes is the preferred method for incorporating transgenes in practicing the invention.

Retroviral infection can also be used to introduce *HDx* transgenes into a non-human animal. The developing non-human embryo can be cultured in vitro to the blastocyst stage. During this time, the blastomeres can be targets for retroviral infection (Jaenich, R. (1976) PNAS 73:1260-1264). Efficient infection of the blastomeres is obtained by enzymatic treatment to remove the zona pellucida (Manipulating the Mouse Embryo, Hogan eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1986). The viral vector system used to introduce the transgene is typicaly a replication-defective retrovirus carrying the transgene (Jahner et al. (1985) PNAS 82:6927-6931; Van der Putten et al. (1985) PNAS 82:6148-6152). Transfection is easily and efficiently obtained by culturing the blastomeres on a monolayer of virus-producing cells (Van der Putten, supra; Stewart et al. (1987) EMBO J. 6: 383-388). Alternatively, infection can be performed at a later stage. Virus or virus-producing cells can be injected into the blastocoele (Jahner et al. (1982) Nature 298: 623-628). Most of the founders will be mosaic for the transgene since incorporation occurs only in a subset of the cells which formed the transgenic non-human animal. Further, the founder may contain various retrovial insertions of the transgene at different positions in the genome which generally will segregate in the offspring. In addition, it is also possible to introduce transgenes into the germ line by intrauterine retroviral infection of the midgestation embryo (Jahner et al. (1982) supra).

A third type of target cell for transgene introduction is the embryonic stem cell (ES). ES cells are obtained from pre-implantation embryos cultured in vitro and fused with embryos (Evans et al (1981) Nature 292: 154-156; Bradley et al. (1984) Nature 309: 255-258; Gossler et al. (1986) PNAS 83: 9065-9069; and Robertson et al (1986) Nature 322: 445-448). Transgenes can be efficiently introduced into the ES cells by DNA transfection or by retrovirus-mediated transduction. Such transformed ES cells can thereafter be combined with blastocysts from a

non-human animal. The ES cells thereafter colonize the embryo and contribute to the germ line of the resulting chimeric animal. For review see Jaenisch, R. (1988) Science 240: 1468-1474.

Methods of making *HDx* knock-out of disruption transgenic animals are also generally known. See, for example, Manipulating the Mouse Embryo, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Recombinase dependent knockouts can also be generated, e.g., by homologous recombination to insert recombinase target sequences flanking portions of an endogenous *HDx* gene, such that tissue specific and/or temporal control of inactivation of an *HDx* allele can be controlled as above.

B. Exemplary Polypeptides

The present invention makes available isolated *HDx* polypeptides which are isolated from, or otherwise substantially free of other cellular proteins. The term "substantially free of other cellular proteins" (also referred to herein as "contaminating proteins") or "substantially pure or purified preparations" are defined as encompassing preparations of *HDx* polypeptides having less than about 20% (by dry weight) contaminating protein, and preferably having less than about 5% contaminating protein. Functional forms of the subject polypeptides can be prepared, for the first time, as purified preparations by using a cloned gene as described herein. Preferred *HDx* proteins of the invention have an amino acid sequence which is at least about 60%, 70%, 80%, 85%, 90%, or 95% identical or homologous to an amino acid sequence of any one of SEQ ID Nos. 1, 3 or 5. Even more preferred *HDx* proteins comprise an amino acid sequence which is at least about 97, 98, or 99% homologous or identical to an amino acid sequence of any one of SEQ ID Nos. 1, 3, or 5. Such proteins can be recombinant proteins, and can be, e.g., produced *in vitro* from nucleic acids comprising a nucleotide sequence set forth in SEQ ID Nos.. 1, 3, or 5, or homologs thereof. For example, recombinant polypeptides preferred by the present invention can be encoded by a nucleic acid, which is at least 85% homologous and more preferably 90% homologous and most preferably 95% homologous with a nucleotide sequence set forth in SEQ ID NOS. 1, 3 or 5. Polypeptides which are encoded by a nucleic acid that is at least about 98-99% homologous with the sequence of SEQ ID NOS: 1, 3 or 5 are also within the scope of the invention.

In a preferred embodiment, an *HDx* protein of the present invention is a mammalian *HDx* protein, and more preferably a human *HDx* protein. In a particularly preferred embodiment an

HDx protein is set forth as SEQ ID No: 2, 4 or 6. In particularly preferred embodiment, an *HDx* protein has an *HDx* bioactivity. It will be understood that certain post-translational modifications, e.g., phosphorylation and the like, can increase the apparent molecular weight of the *HDx* protein relative to the unmodified polypeptide chain.

HDx polypeptides preferably are capable of functioning in one of either role of an agonist or antagonist of at least one biological activity of a wild-type ("authentic") *HDx* protein of the appended sequence listing. The term "evolutionarily related to", with respect to amino acid sequences of *HDx* proteins, refers to both polypeptides having amino acid sequences which have arisen naturally, and also to mutational variants of human *HDx* polypeptides which are derived, for example, by combinatorial mutagenesis. Full length proteins or fragments corresponding to one or more particular motifs and/or domains or to arbitrary sizes, for example, at least 5, 10, 25, 50, 75 and 100, amino acids in length are within the scope of the present invention.

For example, isolated *HDx* polypeptides can be encoded by all or a portion of a nucleic acid sequence shown in any of SEQ ID NOS. 1, 3 or 5. Isolated peptidyl portions of *HDx* proteins can be obtained by screening peptides recombinantly produced from the corresponding fragment of the nucleic acid encoding such peptides. In addition, fragments can be chemically synthesized using techniques known in the art such as conventional Merrifield solid phase f-Moc or t-Boc chemistry. For example, an *HDx* polypeptide of the present invention may be arbitrarily divided into fragments of desired length with no overlap of the fragments, or preferably divided into overlapping fragments of a desired length. The fragments can be produced (recombinantly or by chemical synthesis) and tested to identify those peptidyl fragments which can function as either agonists or antagonists of a wild-type (e.g., "authentic") *HDx* protein.

In general, polypeptides referred to herein as having an activity (e.g., are "bioactive") of an *HDx* protein are defined as polypeptides which include an amino acid sequence encoded by all or a portion of the nucleic acid sequences shown in one of SEQ ID NOS: 1, 3 or 5, and which mimic or antagonize all or a portion of the biological/biochemical activities of a naturally occurring *HDx* protein. Examples of such biological activity include the ability to deacetylate histones, bind to histone(s), 14-3-3 proteins, MEF2 transcription factors and/or RbAp48. Other biological activities of the subject *HDx* proteins are described herein or will be reasonably apparent to those skilled in the art. According to the present invention, a polypeptide has

biological activity if it is a specific agonist or antagonist of a naturally-occurring form of an *HDx* protein.

Other preferred proteins of the invention are those encoded by the nucleic acids set forth in the section pertaining to nucleic acids of the invention. In particular, the invention provides fusion proteins, e.g., *HDx*-immunoglobulin fusion proteins. Such fusion proteins can provide, e.g., enhanced stability and solubility of *HDx* proteins and may thus be useful in therapy.

In addition to utilizing fusion proteins to enhance immunogenicity, it is widely appreciated that fusion proteins can also facilitate the expression of proteins, and accordingly, can be used in the expression of the *HDx* polypeptides of the present invention. For example, *HDx* polypeptides can be generated as glutathione-S-transferase (GST-fusion) proteins. Such GST-fusion proteins can enable easy purification of the *HDx* polypeptide, as for example by the use of glutathione-derivatized matrices (see, for example, Current Protocols in Molecular Biology, eds. Ausubel et al. (N.Y.: John Wiley & Sons, 1991)).

The present invention further pertains to methods of producing the subject *HDx* polypeptides. For example, a host cell transfected with a nucleic acid vector directing expression of a nucleotide sequence encoding the subject polypeptides can be cultured under appropriate conditions to allow expression of the peptide to occur. Suitable media for cell culture are well known in the art. The recombinant *HDx* polypeptide can be isolated from cell culture medium, host cells, or both using techniques known in the art for purifying proteins including ion-exchange chromatography, gel filtration chromatography, ultrafiltration, electrophoresis, and immunoaffinity purification with antibodies specific for such peptide. In a preferred embodiment, the recombinant *HDX* polypeptide is a fusion protein containing a domain which facilitates its purification, such as GST fusion protein.

Moreover, it will be generally appreciated that, under certain circumstances, it may be advantageous to provide homologs of one of the subject *HDX* polypeptides which function in a limited capacity as one of either an *HDX* agonist (mimetic) or an *HDX* antagonist, in order to promote or inhibit only a subset of the biological activities of the naturally-occurring form of the protein. Thus, specific biological effects can be elicited by treatment with a homolog of limited function, and with fewer side effects relative to treatment with agonists or antagonists which are directed to all of the biological activities of naturally occurring forms of *HDX* proteins.

Homologs of each of the subject *HDX* proteins can be generated by mutagenesis, such as by discrete point mutation(s), or by truncation. For instance, mutation can give rise to homologs which retain substantially the same, or merely a subset, of the biological activity of the *HDX* polypeptide from which it was derived. Alternatively, antagonistic forms of the protein can be generated which are able to inhibit the function of the naturally occurring form of the protein, such as by competitively binding to an *HDX* receptor.

The recombinant *HDX* polypeptides of the present invention also include homologs of the wildtype *HDX* proteins, such as versions of those protein which are resistant to proteolytic cleavage, as for example, due to mutations which alter ubiquitination or other enzymatic targeting associated with the protein.

HDX polypeptides may also be chemically modified to create *HDX* derivatives by forming covalent or aggregate conjugates with other chemical moieties, such as glycosyl groups, lipids, phosphate, acetyl groups and the like. Covalent derivatives of *HDX* proteins can be prepared by linking the chemical moieties to functional groups on amino acid sidechains of the protein or at the N-terminus or at the C-terminus of the polypeptide.

Modification of the structure of the subject *HDX* polypeptides can be for such purposes as enhancing therapeutic or prophylactic efficacy, stability (e.g., ex vivo shelf life and resistance to proteolytic degradation), or post-translational modifications (e.g., to alter phosphorylation pattern of protein). Such modified peptides, when designed to retain at least one activity of the naturally-occurring form of the protein, or to produce specific antagonists thereof, are considered functional equivalents of the *HDX* polypeptides described in more detail herein. Such modified peptides can be produced, for instance, by amino acid substitution, deletion, or addition. The substitutional variant may be a substituted conserved amino acid or a substituted non-conserved amino acid.

For example, it is reasonable to expect that an isolated replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, a threonine with a serine, or a similar replacement of an amino acid with a structurally related amino acid (i.e. isosteric and/or isoelectric mutations) will not have a major effect on the biological activity of the resulting molecule. Conservative replacements are those that take place within a family of amino acids that are related in their side chains. Genetically encoded amino acids can be divided into four families: (1) acidic = aspartate, glutamate; (2) basic = lysine, arginine, histidine; (3) nonpolar =

alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan; and (4) uncharged polar = glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine. In similar fashion, the amino acid repertoire can be grouped as (1) acidic = aspartate, glutamate; (2) basic = lysine, arginine histidine, (3) aliphatic = glycine, alanine, valine, leucine, isoleucine, serine, threonine, with serine and threonine optionally be grouped separately as aliphatic-hydroxyl; (4) aromatic = phenylalanine, tyrosine, tryptophan; (5) amide = asparagine, glutamine; and (6) sulfur -containing = cysteine and methionine. (see, for example, Biochemistry, 2nd ed., Ed. by L. Stryer, WH Freeman and Co.: 1981). Whether a change in the amino acid sequence of a peptide results in a functional *HDX* homolog (e.g., functional in the sense that the resulting polypeptide mimics or antagonizes the wild-type form) can be readily determined by assessing the ability of the variant peptide to produce a response in cells in a fashion similar to the wild-type protein, or competitively inhibit such a response. Polypeptides in which more than one replacement has taken place can readily be tested in the same manner.

This invention further contemplates a method for generating sets of combinatorial mutants of the subject *HDX* proteins as well as truncation mutants, and is especially useful for identifying potential variant sequences (e.g., homologs). The purpose of screening such combinatorial libraries is to generate, for example, novel *HDX* homologs which can act as either agonists or antagonist, or alternatively, possess novel activities all together. Thus, combinatorially-derived homologs can be generated to have an increased potency relative to a naturally occurring form of the protein.

In one embodiment, the variegated library of *HDX* variants is generated by combinatorial mutagenesis at the nucleic acid level, and is encoded by a variegated gene library. For instance, a mixture of synthetic oligonucleotides can be enzymatically ligated into gene sequences such that the degenerate set of potential *HDX* sequences are expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage display) containing the set of *HDX* sequences therein.

There are many ways by which such libraries of potential *HDX* homologs can be generated from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be carried out in an automatic DNA synthesizer, and the synthetic genes then ligated into an appropriate expression vector. The purpose of a degenerate set of genes is to provide, in one mixture, all of the sequences encoding the desired set of potential *HDX*

sequences. The synthesis of degenerate oligonucleotides is well known in the art (see for example, Narang, SA (1983) Tetrahedron 39:3; Itakura et al. (1981) Recombinant DNA, Proc 3rd Cleveland Sympos. Macromolecules, ed. AG Walton, Amsterdam: Elsevier pp 273-289; Itakura et al. (1984) Annu. Rev. Biochem. 53:323; Itakura et al. (1984) Science 198:1056; Ike et al. (1983) Nucleic Acid Res. 11:477. Such techniques have been employed in the directed evolution of other proteins (see, for example, Scott et al. (1990) Science 249:386-390; Roberts et al. (1992) PNAS 89:2429-2433; Devlin et al. (1990) Science 249: 404-406; Cwirla et al. (1990) PNAS 87: 6378-6382; as well as U.S. Patents NOS. 5,223,409, 5,198,346, and 5,096,815).

Likewise, a library of coding sequence fragments can be provided for an *HDX* clone in order to generate a variegated population of *HDX* fragments for screening and subsequent selection of bioactive fragments. A variety of techniques are known in the art for generating such libraries, including chemical synthesis. In one embodiment, a library of coding sequence fragments can be generated by (i) treating a double stranded PCR fragment of an *HDX* coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule; (ii) denaturing the double stranded DNA; (iii) renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products; (iv) removing single stranded portions from reformed duplexes by treatment with S1 nuclease; and (v) ligating the resulting fragment library into an expression vector. By this exemplary method, an expression library can be derived which codes for N-terminal, C-terminal and internal fragments of various sizes.

A wide range of techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a certain property. Such techniques will be generally adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of *HDX* homologs. The most widely used techniques for screening large gene libraries typically comprises cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates relatively easy isolation of the vector encoding the gene whose product was detected. Each of the illustrative assays described below are amenable to high through-put analysis as necessary to screen large numbers of degenerate *HDX* sequences created by combinatorial mutagenesis techniques. Combinatorial

mutagenesis has a potential to generate very large libraries of mutant proteins, e.g., in the order of 10^{26} molecules. Combinatorial libraries of this size may be technically challenging to screen even with high throughput screening assays. To overcome this problem, a new technique has been developed recently, recrusive ensemble mutagenesis (REM), which allows one to avoid the very high proportion of non-functional proteins in a random library and simply enhances the frequency of functional proteins, thus decreasing the complexity required to achieve a useful sampling of sequence space. REM is an algorithm which enhances the frequency of functional mutants in a library when an appropriate selection or screening method is employed (Arkin and Yourvan, 1992, PNAS USA 89:7811-7815; Yourvan et al., 1992, Parallel Problem Solving from Nature, 2., In Maenner and Manderick, eds., Elsevir Publishing Co., Amsterdam, pp. 401-410; Delgrave et al., 1993, Protein Engineering 6(3):327-331).

The invention also provides for reduction of the *HDX* proteins to generate mimetics, e.g., peptide or non-pepide agents, such as small molecules, which are able to disrupt binding of an *HDX* polypeptide of the present invention with a molecule, e.g. target peptide. Thus, such mutagenic techniques as described above are also useful to map the determinants of the *HDX* proteins which participate in protein-protein interactions involved in, for example, binding of the subject *HDX* polypeptide to a target peptide. To illustrate, the critical residues of a subject *HDX* polypeptide which are involved in molecular recognition of its receptor can be determined and used to generate *HDX* derived peptidomimetics or small molecules which competitively inhibit binding of the authentic *HDX* protein with that moiety. By employing, for example, scanning mutagenesis to map the amino acid residues of the subject *HDX* proteins which are involved in binding other proteins, peptidomimetic compounds can be generated which mimic those residues of the *HDX* protein which facilitate the interaction. Such mimetics may then be used to interfere with the normal function of an *HDX* protein. For instance, non-hydrolyzable peptide analogs of such residues can be generated using benzodiazepine (e.g., see Freidinger et al. in Peptides: Chemistry and Biology, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), azepine (e.g., see Huffman et al. in Peptides: Chemistry and Biology, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), substituted gamma lactam rings (Garvey et al. in Peptides: Chemistry and Biology, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), keto-methylene pseudopeptides (Ewenson et al. (1986) J Med Chem 29:295; and Ewenson et al. in Peptides: Structure and Function (Proceedings of the 9th American Peptide Symposium) Pierce Chemical Co. Rockland, IL, 1985), b-turn dipeptide cores (Nagai et al. (1985) Tetrahedron Lett 26:647; and Sato et al. (1986) J Chem Soc Perkin Trans 1:1231), and b-aminoalcohols (Gordon et al. (1985) Biochem Biophys Res Commun126:419; and Dann et al. (1986) Biochem Biophys Res Commun 134:71).

C. *HDx* Antibodies

Another aspect of the invention pertains to an antibody specifically reactive with an *HDx* protein. For example, by using immunogens derived from an *HDx* protein, e.g., based on the cDNA sequences, anti-protein/anti-peptide antisera or monoclonal

antibodies can be made by standard protocols (See, for example, Antibodies: A Laboratory Manual ed. By Halow and Lane (Cold Spring Harbor Press: 1988)). A mammal, such as a mouse, a hamster or rabbit can be immunized with an immunogenic form of the peptide (e.g., an *HDx* polypeptide or an antigenic fragment which is capable of eliciting an antibody response). Techniques for conferring immunogenicity on a protein or peptide include conjugation to carriers or other techniques well known in the art. An immunogenic portion of an *HDx* protein can be administered in the presence of adjuvant. The progress of immunization can be monitored by detection of antibody titers in plasma or serum. Standard ELISA or other immunoassays can be used with the immunogen as antigen to assess the levels of antibodies. In a preferred embodiment, the subject antibodies are immunospecific for antigenic determinants of a class II *HDx* protein of an organism, such as a mammal, e.g, antigenic determinants of a protein represented by one of SEQ ID Nos. or closely related homologs (e.g., at least 85% homologous, preferably at least 90% homologous, and more preferably at least 90% identical). In yet a further embodiment of the present invention, in order to provide, for example, antibodies which are immuno-selective for discrete *HDx* homologs, e.g, HDAC 4, HDAC5 or HDAC6, the anti-*HDx* polypeptide antibodies do not substantially cross react (i.e., does not react specifically) with a protein which is, for example, less than 85%, 90% or 95% homologous with the selected *HDx*. By "not substantially cross react", it is meant that the antibody has a binding affinity for a non-homologous protein which is at least one order of magnitude, more preferably at least 2 orders of magnitude, and even more preferably at least 3 orders of magnitude less than the binding affinity of the antibody for the intended target *HDx*.

Following immunization of an animal with an antigenic preparation of an *HDx* polypeptide, anti-*HDx* antisera can be obtained and, if desired, polyclonal anti-*HDx* antibodies can be isolated from the serum. To produce monoclonal antibodies, antibody-producing cells (lymphocytes) can be harvested from an immunized animal and fused by standard somatic cell fusion procedures with immortalizing cells such as myeloma cells to yield hybridoma cells. Such techniques are well known in the art, and include, for example, the hybridoma technique (originally developed by Kohler and Milstein, (1975) Nature, 256: 495-497), the human B cell hybridoma technique (Kozbar et al., (1983)

Immunology Today, 4: 72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al., (1985) Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc. pp. 77-96). Hybridoma cells can be screened immunochemically for production of antibodies specifically reactive with an *HDx* polypeptide of the present invention and monoclonal antibodies isolated from a culture comprising such hybridoma cells.

The term antibody, as used herein, is intended to include fragments thereof which are also specifically reactive with one of the subject *HDx* polypeptides. Antibodies can be fragmented using conventional techniques and the fragments screened for utility in the same manner as described above for whole antibodies. For example F(ab)2 fragments can be generated by treating antibody with pepsin. The resulting F(ab)2 fragment can be treated to reduce disulfide bridges to produce Fab fragments. The antibody of the present invention is further intended to include bispecific and chimeric molecules having affinity for an *HDx* protein conferred by at least one CDR region of the antibody.

Both monoclonal and polyclonal antibodies (Ab) directed against authentic *HDx* polypeptides, or *HDx* variants, and antibody fragments such as Fab, F(ab)2, Fv and scFv can be used to block the action of one or more *HDx* proteins and allow the study of the role of these proteins in, for example, differentiation of tissue. Experiments of this nature can aid in deciphering the role of *HDx* proteins that may be involved in control of proliferation versus differentiation, e.g., in patterning and tissue formation.

Antibodies which specifically bind *HDx* epitopes can also be used in immunohistochemical staining of tissue samples in order to evaluate the abundance and pattern of expression of each of the subject *HDx* polypeptides. Anti-*HDx* antibodies can be used diagnostically in immuno-precipitation and immuno-blotting to detect and evaluate *HDx* protein levels in tissue as part of a clinical testing procedure. For instance, such measurements can be useful in predictive valuations of the onset or progression of proliferative or differentiative disorders. Likewise, the ability to monitor *HDx* protein levels in an individual can allow determination of the efficacy of a given treatment regimen for an individual affected with such a disorder. The level of *HDx* polypeptides may be measured from cells in bodily fluid, such as in samples of cerebral spinal fluid or amniotic fluid, or can be measured in tissue, such as produced by biopsy. Diagnostic

assays using anti-*HDx* antibodies can include, for example, immunoassays designed to aid in early diagnosis of a disorder, particularly ones which are manifest at birth. Diagnostic assays using anti-*HDx* polypeptide antibodies can also include immunoassays designed to aid in early diagnosis and phenotyping neoplastic or hyperplastic disorders.

D. *HDx* Therapeutic Agents

As discussed above, purified and recombinant *HDx* polypeptides are made available by the present invention and thus facilitates the development of assays which can be used to screen for drugs, including *HDx* homologs, which are either agonists of antagonists of the normal cellular function of the subject *HDx* polypeptides, or of their role in the pathogenesis of cellular differentiation and/or proliferation and disorders related thereto. In certain embodiments, the subject method is used to identify agents which potentiate or inhibit the deacetylase activity of an *HDx* protein. Moreover, because proteins have been identified which bind to the subject *HDx* proteins, e.g., such as histones, 14-3-3 proteins, MEF2 transcription factor, and RbAp48as, the present invention further provides drug screening assays for detecting agents which modulate those interactions.

In a general sense, the assay evaluates the ability of a compound to modulate binding between an *HDx* polypeptide and a molecule, be it protein or DNA, that interacts with the *HDx* polypeptide, be it a substrate of the deacteylase, or serves a regulatory function. Exemplary compounds which can be screened include peptides, nucleic acids, carbohydrates, small organic molecules, and natural product extract libraries, such as isolated from animals, plants, fungus and/or microbes.

It is contemplated that any of the novel interactions described herein could be exploited in a drug screening assay. To illustrate, the interaction between an *HDx* protein and a histone, a 14-3-3 protein or other *HDx*-binding protein can be detected in the presence and the absence of a test compound. In other embodiments, the ability of a compound to modulate the acetylase activity of an *HDx* protein can be assessed. A variety of assay formats will suffice and, in light of the present inventions, will be comprehended by a skilled artisan.

In a preferred embodiment, assays which employ the subject mammalian *HDx* proteins can be used to identify compounds that have therapeutic indexes more favorable than sodium butyrate, trapoxin, trichostatin or the like. For instance, trapoxin-like drugs can be identified by the present invention which have enhanced tissue-type or cell-type specificity relative to trapoxin. To illustrate, the subject assays can be used to generate compounds which preferentially inhibit IL-2 mediated proliferation/activation of lymphocytes, or inhibit proliferation of certain tumor cells, without substantially interfering with other tissues, e.g. hepatocytes. Likewise, similar assays can be used to identify drugs which inhibit proliferation of yeast cells or other lower eukaryotes, but which have a substantially reduced effect on mammalian cells, thereby improving therapeutic index of the drug as an anti-mycotic agent.

In one embodiment, the identification of such compounds is made possible by the use of differential screening assays which detect and compare drug-mediated inhibition of deacetylase activity or protein-protein or protein-DNA interactions involving two or more different *HDx* enzymes, e.g., to find compounds that selectively inhibit class I or class I *HDx* proteins relative to one another or selectively inhibit one *HDx* protein relative to all the others. To illustrate, the assay can be designed for side-by-side comparison of the effect of a test compound on the deacetylase activity or protein interactions of tissue-type specific *HDx* proteins. Given the apparent diversity of *HDx* proteins, it is probable that different functional *HDx* activities, or *HDx* complexes exist and, in certain instances, are localized to particular tissue or cell types. Thus, test compounds can be screened for agents able to inhibit the tissue-specific formation of only a subset of the possible repetoire of *HDx*/regulatory protein complexes, or which preferentially inhibit certain *HDx* enzymes. In an exemplary embodiment, an interaction trap assay can be derived using class I and class II *HDx* "bait" proteins, while the "fish" protein is constant in each, e.g., a human RbAp48 construct. Running the interaction trap side-by-side permits the detection for agents which have a greater effect (e.g., statistically significant) on the formation of one of the class I *HDx*/RbAp48 complexes than on the formation of the class I *HDx*/RbAp48 complexes.

In similar fashion, differential screening assays can be used to exploit the difference in protein interactions and/or catalytic mechanism of mammalian *HDx* proteins

and yeast *RPD3* proteins in order to identify agents which display a statistically significant increase in specificity for inhibiting the yeast enzyme relative to the mammalian enzyme. Thus, lead compounds which act specifically on pathogens, such as fungus involved in mycotic infections, can be developed. By way of illustration, the present assays can be used to screen for agents which may ultimately be useful for inhibiting at least one fungus implicated in such mycosis as candidiasis, aspergillosis, mucormycosis, blastomycosis, geotrichosis, cryptococcosis, chromoblastomycosis, coccidioidomycosis, conidiosporosis, histoplasmosis, maduromycosis, rhinosporidosis, nocaidiosis, para-actinomycosis, penicilliosis, monoliasis, or sporotrichosis. For example, if the mycotic infection to which treatment is desired is candidiasis, the present assay can comprise comparing the relative effectiveness of a test compound on inhibiting the deacetylase activity of a mammalian *HDx* protein with its effectiveness towards inhibiting the deacetylase activity of an *RPD3* homolog cloned from yeast selected from the group consisting of *Candida albicans*, *Candida stellatoidea*, *Candida tropicalis*, *Candida parapsilosis*, *Candida krusei*, *Candida pseudotropicalis*, *Candida quillermondii*, or *Candida rugosa*. Likewise, the present assay can be used to identify anti-fungal agents which may have therapeutic value in the treatment of aspergillosis by selectively targeting *RPD3* homologs cloned from yeast such as *Aspergillus fumigatus*, *Aspergillus flavus*, *Aspergillus niger*, *Aspergillus nidulans*, or *Aspergillus terreus*. Where the mycotic infection is mucormycosis, the *RPD3* deacetylase can be derived from yeast such as *Rhizopus arrhizus*, *Rhizopus oryzae*, *Absidia corymbifera*, *Absidia ramosa*, or *Mucor pusillus*. Sources of other *RPD3* activities for comparison with a mammalian *HDx* activity includes the pathogen *Pneumocystis carinii*.

In addition to such *HDx* therpeutic uses, anti-fungal agents developed with such differential screening assays can be used, for example, as preservatives in foodstuff, feed supplement for promoting weight gain in livestock, or in disinfectant formulations for treatment of non-living matter, e.g., for decontaminating hospital equipment and rooms.

In a similar fashion, side by side comparison of inhibition of mammalian *HDx* proteins and an insect *HDx*-related proteins, will permit selection of *HDx* inhibitors which discriminate between the human/mammalian and insect enzymes. Accordingly,

the present invention expressly contemplates the use and formulations of the subject *HDx* therpeutics in insecticides, such as for use in management of insects like the fruit fly.

In yet another embodiment, certain of the subject *HDx* inhibitors can be selected on the basis of inhibitory specificity for plant *HDx*-related activities relative to the mammalian enzyme. For example, a plant *HDx*-related protein can be disposed in a differential screen with one or more of the human enzymes to select those compounds of greatest selectivity for inhibiting the plant enzyme. Thus, the present invention specifically contemplates formaulations of the subject *HDx* inhibitors for agricultural applications, such as in the form of a defoliant or the like.

In many drug screening programs which test libraries of compounds and natural products, high throuput assays are desirable in order to maximize the number of compounds surveyed in a given period of time. Assays which are performed in cell-free systems, such as may be derived with purified or semi-purified proteins, are often preferred as "primary" screens in that they can be generated to permit rapid development and relatively easy detection of an alteration in a molecular target which is mediated by a test compound. Moreover, the effects of cellular toxicity and/or bioavailablity of the test copound can be generally ignored in the in vitro system, the assay instead being focused primarily on the effect of the drug on the molecular target as may be manifest in an alteration of binding affinity with upstream or downstream elements. Accordingly, in an exemplary screening assay of the present invention, a reaction mixture is generated to include an *HDx* polypeptide, compound(s) of interest, and a "target polypeptide" e.g., a protein which interacts with the *HDx* polypeptide, whether as a substrate or by some other protein-protein interaction. Detection and quantification of complexes containing the *HDx* protein provide a means for determining a compound's efficacy at inhibiting (or potentiating) complex formation between the *HDx* and the target polypeptide. The efficacy of the compound can be assessed by generating dose response curves from data obtained using various concentrations of the test compound. Moreover, a control assay can also be performed to provide a baseline for comparison. In the control assay, isolate and purified *HDx* polypeptide is added to a composition containing the target polypeptide and the formation of a complex is quantitated in the absence of the test compound.

Complex formation between the *HDx* polypeptide and the target polypeptide may be detected by a variety of techniques. Modulation of the formation of complexes can be quantitated using, for example, detectably labeled proteins such as radiolabeled, fluorescently labeled, or enzymaticaly labeled *HDx* polypeptides, by immunoassay, by chromatographic detection, or by detecting the intrinsic activity of the acetylase.

Typically, it will be desirable to immobilize either *HDx* or the target polypeptide to facilitate separation of complexes from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of *HDx* to the target polypeptide, in the presence and absence of a candidate agent, can be accomplished in any vessel suitable for containing the reactants. Examples include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows the protein to be bound to the matrix. For example, glutathione-S-transferase/*HDx* (GST/*HDx*) fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the cell lysates, e.g. an ^{35}S-labeled, and the test compound, and the mixture incubated under conditions conducive to complex formation, e.g. at physiological conditions for salt and pH, though slightly more stringent conditions may be desired. Following incubation, the beads are washed to remove any unbound label, and the matrix immobilized and radiolabel determined directly (e.g. beads placed in scintillant), or in the supernatant after the complexes are subsequently dissociated. Alternatively, the complexes can be dissociated from the matrix, separated by SDS-PAGE, and the level of *HDx*-binding protein found in the bead fraction quantitated from the gel using standard electrophoretic techniques such as described in the appended examples.

Other techniques for immobilizing proteins on matrices are also available for use in the subject assay. For instance, either *HDx* or target polypeptide can be immobilized utilizing conjugation of biotin and streptavidin. For instance, biotinylated *HDx* molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with *HDx*, but which do not interfere with the

interaction between the *HDx* and target polypeptide, can be derivatized to the wells of the plate, and *HDx* trapped in the wells by antibody conjugation. As above, preparations of an target polypeptide and a test compound are incubated in the *HDx*-presenting wells of the plate, and the amount of complex trapped in the well can be quantitated. Exemplary methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the target polypeptide, or which are reactive with *HDx* protein and compete with the target polypeptide; as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the target polypeptide, either intrinsic or extrinsic activity. In the instance of the latter, the enzyme can be chemically conjugated or provided as a fusion protein with the target polypeptide. To illustrate, the target polypeptide can be chemically cross-linked or genetically fused with horseradish peroxidase, and the amount of polypeptide trapped in the complex can be assessed with a chromogenic substrate of the enzyme, e.g. 3,3'-diamino-benzadine terahydrochloride or 4-chloro-1-napthol. Likewise, a fusion protein comprising the polypeptide and glutathione-S-transferase can be provided, and complex formation quantitated by detecting the GST activity using 1-chloro-2,4-dinitrobenzene (Habig et al (1974) J Biol Chem 249:7130).

For processes which rely on immunodetection for quantitating one of the proteins trapped in the complex, antibodies against the protein, such as anti-*HDx* antibodies, can be used. Alternatively, the protein to be detected in the complex can be "epitope tagged" in the form of a fusion protein which includes, in addition to the *HDx* sequence, a second polypeptide for which antibodies are readily available (e.g. from commercial sources). For instance, the GST fusion proteins described above can also be used for quantification of binding using antibodies against the GST moiety. Other useful epitope tags include myc-epitopes (e.g., see Ellison et al. (1991) J Biol Chem 266:21150-21157) which includes a 10-residue sequence from c-myc, as well as the pFLAG system (International Biotechnologies, Inc.) or the pEZZ-protein A system (Pharamacia, NJ).

In another embodiment of a drug screening, a two hybrid assay can be generated with an *HDx* and *HDx*-binding protein. Drug dependent inhibition or potentiation of the interaction can be scored.

Where the *HDx* proteins themselves, or in complexes with other proteins, are capable of binding DNA and modifying transcription of a gene, a transcriptional based assay using, for example, an transcriptional regulatory sequences responsive to *HDx* complexes operably linked to a detectable marker gene.

Furthermore, each of the assay systems set out above can be generated in a "differential" format as set forth above. That is, the assay format can provide information regarding specificity as well as potency. For instance, side-by-side comparison of a test compound's effect on different *HDx*s can provide information on selectivity, and permit the identification of compounds which selectively modulate the bioactivity of only a subset of the *HDx* family.

Furthermore, inhibitors of the enzymatic activity of each of the subject *HDx* proteins can be identified using assays derived from measuring the ability of an agent to inhibit catalytic conversion of a substrate by the subject proteins. For example, the ability of the subject *HDx* proteins to deacetylate a histone substrate, such as histone H4, in the presence and absence of a candidate inhibitor, can be determined using standard enzymatic assays.

A number of methods have been employed in the art for assaying histone deacetylase activity, and can be incorporated in the drug screening assays of the present invention. In preferred embodiments, the assay will employ a labeled acetyl group linked to appropriate histone lysine residues as substrates. In other embodiments, a histone substrate peptide can be labeled with a group whose signal is dependent on the simultaneous presence or absence of an acetyl group, e.g., the label can be a fluorogenic group whose fluorescence is modulated (either quenched or potentiated) by the presence of the acetyl moiety. Using standard enzymatic analysis, the ability of a test agent to cause a statistically significant change in substrate conversion by a histone deacetylase can be measured, and as desirable, inhibition constants, e.g., K_i values, can be calculated. The histone substrate can be provided as a purified or semi-purified polypeptide or as part of a cell lysate. Likewise, the histone deacetylase can be provided to the reaction mixture as a purified or semi-purified polypeptide or as a cell lysate. Accordingly, the reaction mixtures of the subject method can range from reconstituted protein mixtures derived

with purified preparations of histones and deacetylases, to mixtures of cell lysates, e.g., by admixing baculovirus lysates containing recombinant histones and deacetylases.

In an exemplary embodiment, the histone substrate for the subject assay is provided by isolation of radiolabeled histones from metabolically labelled cells. To illustrate, as described by Hay et al. (1983) *J Biol Chem* 258:3726-3734, HeLa cells can be labelled in culture by addition of [^{3}H]acetate (New England Nuclear) to the culture media. The addition of butyrate, trapoxin or the like can be used to increase the abundance of acetylated histones in the cells. Radiolabelled histones can be isolated from the cells by extraction with H_SSO_4 (Marushige et al. (1966) *J Mol Biol* 15:160-174). Briefly, cells are homogenized in buffer, centrifuged to isolate a nuclear pellet, the subsequently homogenized nuclear pellet centrifuged through sucrose, and the resulting chromatin pellet extracted by addition of H_SSO_4 to yield [^{3}H]acetyl-labelled histones. In an alternate embodiment, nucleosome preparations containing [^{3}H]acetyl-labelled histones can be isolated from the labelled cells. As described in the art, nucleosomes can be isolated from cell preparations by sucrose gradient centrifugation (Hay et al. (1983) *J Biol Chem* 258:3726-3734; and Noll (1967) *Nature* 215:360-363), and polynucleosomes can be prepared by NaCl precipitation from micrococcal nuclease digested cells (Hay et al., *supra*). Similar procedures for isolating labelled histones from other cells types, including yeast, have been described. See, for example, Alonso et al. (1986) *Biochem Biophys Acta* 866:161-169; and Kreiger et al. (1974) *J Biol Chem* 249:332-334. In yet other embodiments, the histone is generated by recombinant gene expression, and includes an exogenous tag (e.g., an HA epitope, a poly(his) sequence or the like) which facilitates in purification from cell extracts. In still other embodiments, whole nuclei can be isolated from metabolically labelled cells by micrococcal nuclease digestion (Hay et al., *supra*)

In still another embodiment, the deacetylase substrate can be provided as an acetylated peptide including a sequence corresponding to the sequence about the specific lysyl residues acetylated on histone, e.g., a peptidyl portions of the core histones H2A, H2B, H3 or H4. Such fragments can be produced by cleavage of acetylated histones derived from metabolically labelled cells, e.g., such as by treatment with proteolytic

enzymes or cyanogen bromide (Kreiger et al., *supra*). In other embodiments, the acetylated peptide can be provided by standard solid phase synthesis using acetylated lysine residues (Kreiger et al., *supra*).

Continuing with the illustrative use of [^{3}H]acetyl-labelled histones, the activity of a histone deacetylase in the subject assays is detected by measuring release of [^{3}H]acetate by standard scintillant techniques. In a merely illustrative example, a reaction mixture is provided which comprises a recombinant *HDx* protein suspended in buffer, along with a sample of [^{3}H]acetyl-labelled histones and (optionally) a test compound. The reaction mixture is maintained at a desired temperature and pH, such as 22°C at pH7.8, for several hours, and the reaction terminated by boiling or other form of denaturation. Released [^{3}H]acetate is extracted and counted. For example, the quenched reaction mixture can be acidified with concentrated HCl, and used to create a biphasic mixture with ethyl acetate. The resulting 2 phase system is thoroughly mixed, centrifuged, and the ethyl acetate phase collected and counted by standard scintillation methods. Other methods for detecting acetate release will be easily recognized by those skilled in the art.

In yet another embodiment, the drug screening assay is derived to include a whole cell recombinantly expressing one or more of a target protein or *HDx* protein. The ability of a test agent to alter the activity of the *HDx* protein can be detected by analysis of the recombinant cell. For example, agonists and antagonists of the *HDx* biological activity can by detected by scoring for alterations in growth or differentiation (phenotype) of the cell. General techniques for detecting each are well known, and will vary with respect to the source of the particular reagent cell utilized in any given assay.

For example, quantification of proliferation of cells in the presence and absence of a candidate agent can be measured with a number of techniques well known in the art, including simple measurement of population growth curves. For instance, where the assay involves proliferation in a liquid medium, turbidimetric techniques (i.e. absorbence/ transmittance of light of a given wavelength through the sample) can be utilized. For example, in the instance where the reagent cell is a yeast cell, measurement of absorbence of light at a wavelength between 540 and 600nm can provide a conveniently fast measure of cell growth. Likewise, ability to form colonies in solid medium (e.g. agar) can be used

to readily score for proliferation. In other embodiments, an *HDx* substrate protein, such as a histone, can be provided as a fusion protein which permits the substrate to be isolated from cell lysates and the degree of acetylation detected. Each of these techniques are suitable for high through-put analysis necessary for rapid screening of large numbers of candidate agents.

In addition, where the ability of an agent to cause or reverse a transformed phenotype, growth in solid media such as agar can further aid in establishing whether a mammalian cell is transformed.

Additionally, visual inspection of the morphology of the reagent cell can be used to determine whether the biological activity of the targeted *HDx* protein has been affected by the added agent. To illustrate, the ability of an agent to influence an apoptotic phenotype which is mediated in some way by a recombinant *HDx* protein can be assessed by visual microscopy. Likewise, the formation of certain cellular structures as part of differentiation, such as the formation of neuritic process, can be visualized under a light microscope.

The nature of the effect of test agent on reagent cell can be assessed by measuring levels of expression of specific genes, e.g., by reverse transcription-PCR. Another method of scoring for effect on *Hdx* activity is by detecting cell-type specific marker expression through immunofluorescent staining. Many such markers are known in the art, and antibodies are readily available. For example, the presence of chondroitin sulphate proteoglycans as well as type-II collagen are correlated with cartilage production in chondrocytes, and each can be detected by immunostaining. Similarly, the human kidney differentiation antigen gp160, human aminopeptidase A, is a marker of kidney induction, and the cytoskeletal protein troponin I is a marker of heart induction. In yet another embodiment, the alteration of expression of a reporter gene construct provided in the reagent cell provides a means of detecting the effect on *HDx* activity. For example, reporter gene constructs derived using the transcriptional regulatory sequences, e.g. the promoters, for developmentally regulated genes can be used to drive the expression of a detectable marker, such as a luciferase gene. In an illustrative embodiment, the construct is derived using the promoter sequence from a gene expressed in a particular differentiative phenotype.

It is also deemed to be within the scope of this invention that the recombinant *HDx* cells of the present assay can be generated so as to comprise heterologous *HDx* proteins (i.e. cross-species expression). For example, *HDx* proteins from one species can be expressed in the cells of another under conditions wherein the heterologous protein is able to rescue loss-of-function mutations in the host cell. For example, the reagent cell can be a yeast cell in which a human *MDx* protein (e.g. exogenously expressed) is the intended target for development of an anti-proliferative agent. To illustrate, the M778 strain, *MATa ura3-52 trp1 ι his3-200 leu2-1 trk1 rpd3 ::HIS3*, described by Vidal et al. (1991) *Mol Cell Biol* 6317-6327, which lacks a functional endogenous RPD3 gene can be transfected with an expression plasmid including a mammalian *HDx* gene in order to complement the RPD3 loss-of-function. For example, the coding sequence for *HD4*, *HD5* or *HD6* can be cloned into a pRS integrative plasmid containing a selectable marker (Sikorski et al. (1989) *Genetics* 122:19-27), and resulting construct used to transform the M778 strain. The resulting cells should produce a mammalian *HDx* protein which may be capable performing at least some of the functions of the yeast RPD3 protein. The *HDx* transformed yeast cells can be easier to manipulate than mammalian cells, and can provide access to certain assay formats, such as turbidity detection methods, which may not be obtainable with mammalian cells.

Moreover, the combination of the "mammalianized" strain with the strain M537 (*MATa ura3-52 trp1 ι his3-200 leu2-1 trk1* , Vidal et al., *supra*) can provide an exquisitely sensitive cell-based assay for detecting agent which specifically inhibit, for example, the yeast RPD3 deacetylase.

E. *Small Molecule Combinatorial Libraries for HDx Inhibitors*

The subject reactions readily lend themselves to the creation of combinatorial libraries of compounds for the screening of pharmaceutical, agrochemical or other biological or medically-related activity or material-related qualities. A combinatorial library for the purposes of the present invention is a mixture of chemically related compounds which may be screened together for a desired property; said libraries may be in solution or covalently linked to a solid support. The preparation of many related compounds in a single reaction greatly reduces and simplifies the number of screening processes which need to be carried out. Screening for the appropriate biological,

pharmaceutical, agrochemical or physical property may be done by conventional methods.

Diversity in a library can be created at a variety of different levels. For instance, the substrate aryl groups used in a combinatorial approach can be diverse in terms of the core aryl moiety, e.g., a variegation in terms of the ring structure, and/or can be varied with respect to the other substituents.

A variety of techniques are available in the art for generating combinatorial libraries of small organic molecules. See, for example, Blondelle et al. (1995) Trends Anal. Chem. 14:83; the Affymax U.S. Patents 5,359,115 and 5,362,899: the Ellman U.S. Patent 5,288,514: the Still et al. PCT publication WO 94/08051; Chen et al. (1994) JACS 116:2661: Kerr et al. (1993) JACS 115:252; PCT publications WO92/10092, WO93/09668 and WO91/07087; and the Lerner et al. PCT publication WO93/20242). Accordingly, a variety of libraries on the order of about 16 to 1,000,000 or more diversomers can be synthesized and screened for a particular activity or property.

In an exemplary embodiment, a library of substituted diversomers can be synthesized using the subject reactions adapted to the techniques described in the Still et al. PCT publication WO 94/08051, e.g., being linked to a polymer bead by a hydrolyzable or photolyzable group, e.g., located at one of the positions of substrate. According to the Still et al. technique, the library is synthesized on a set of beads, each bead including a set of tags identifying the particular diversomer on that bead. In one embodiment, which is particularly suitable for discovering enzyme inhibitors, the beads can be dispersed on the surface of a permeable membrane, and the diversomers released from the beads by lysis of the bead linker. The diversomer from each bead will diffuse across the membrane to an assay zone, where it will interact with an enzyme assay. Detailed descriptions of a number of combinatorial methodologies are provided below.

Direct Characterization

A growing trend in the field of combinatorial chemistry is to exploit the sensitivity of techniques such as mass spectrometry (MS), e.g., which can be used to characterize sub-femtomolar amounts of a compound, and to directly determine the chemical constitution of a compound selected from a combinatorial library. For instance, where the library is provided on an insoluble support matrix, discrete populations of

compounds can be first released from the support and characterized by MS. In other embodiments, as part of the MS sample preparation technique, such MS techniques as MALDI can be used to release a compound from the matrix, particularly where a labile bond is used originally to tether the compound to the matrix. For instance, a bead selected from a library can be irradiated in a MALDI step in order to release the diversomer from the matrix, and ionize the diversomer for MS analysis.

Multipin Synthesis

The libraries of the subject method can take the multipin library format. Briefly, Geysen and co-workers (Geysen et al. (1984) PNAS 81:3998-4002) introduced a method for generating compound libraries by a parallel synthesis on polyacrylic acid-grated polyethylene pins arrayed in the microtitre plate format. The Geysen technique can be used to synthesize and screen thousands of compounds per week using the multipin method, and the tethered compounds may be reused in many assays. Appropriate linker moieties can also been appended to the pins so that the compounds may be cleaved from the supports after synthesis for assessment of purity and further evaluation (c.f., Bray et al. (1990) Tetrahedron Lett 31:5811-5814; Valerio et al. (1991) Anal Biochem 197:168-177; Bray et al. (1991) Tetrahedron Lett 32:6163-6166).

Divide-Couple-Recombine

In yet another embodiment, a variegated library of compounds can be provided on a set of beads utilizing the strategy of divide-couple-recombine (see, e.g., Houghten (1985) PNAS 82:5131-5135; and U.S. Patents 4,631,211; 5,440,016; 5,480,971). Briefly, as the name implies, at each synthesis step where degeneracy is introduced into the library, the beads are divided into separate groups equal to the number of different substituents to be added at a particular position in the library, the different substituents coupled in separate reactions, and the beads recombined into one pool for the next iteration.

In one embodiment, the divide-couple-recombine strategy can be carried out using an analogous approach to the so-called "tea bag" method first developed by Houghten, where compound synthesis occurs on resin sealed inside porous polypropylene bags (Houghten et al. (1986) PNAS 82:5131-5135). Substituents are coupled to the compound-bearing resins by placing the bags in appropriate reaction solutions, while all

common steps such as resin washing and deprotection are performed simultaneously in one reaction vessel. At the end of the synthesis, each bag contains a single compound.

Combinatorial Libraries by Light-Directed, Spatially Addressable Parallel Chemical Synthesis

A scheme of combinatorial synthesis in which the identity of a compound is given by its locations on a synthesis substrate is termed a spatially-addressable synthesis. In one embodiment, the combinatorial process is carried out by controlling the addition of a chemical reagent to specific locations on a solid support (Dower et al. (1991) Annu Rep Med Chem 26:271-280; Fodor, S.P.A. (1991) Science 251:767; Pirrung et al. (1992) U.S. Patent No. 5,143,854; Jacobs et al. (1994) Trends Biotechnol 12:19-26). The spatial resolution of photolithography affords miniaturization. This technique can be carried out through the use protection/deprotection reactions with photolabile protecting groups.

The key points of this technology are illustrated in Gallop et al. (1994) J Med Chem 37:1233-1251. A synthesis substrate is prepared for coupling through the covalent attachment of photolabile nitroveratryloxycarbonyl (NVOC) protected amino linkers or other photolabile linkers. Light is used to selectively activate a specified region of the synthesis support for coupling. Removal of the photolabile protecting groups by light (deprotection) results in activation of selected areas. After activation, the first of a set of amino acid analogs, each bearing a photolabile protecting group on the amino terminus, is exposed to the entire surface. Coupling only occurs in regions that were addressed by light in the preceding step. The reaction is stopped, the plates washed, and the substrate is again illuminated through a second mask, activating a different region for reaction with a second protected building block. The pattern of masks and the sequence of reactants define the products and their locations. Since this process utilizes photolithography techniques, the number of compounds that can be synthesized is limited only by the number of synthesis sites that can be addressed with appropriate resolution. The position of each compound is precisely known; hence, its interactions with other molecules can be directly assessed.

In a light-directed chemical synthesis, the products depend on the pattern of illumination and on the order of addition of reactants. By varying the lithographic

patterns, many different sets of test compounds can be synthesized simultaneously; this characteristic leads to the generation of many different masking strategies.

Encoded Combinatorial Libraries

In yet another embodiment, the subject method utilizes a compound library provided with an encoded tagging system. A recent improvement in the identification of active compounds from combinatorial libraries employs chemical indexing systems using tags that uniquely encode the reaction steps a given bead has undergone and, by inference, the structure it carries. Conceptually, this approach mimics phage display libraries, where activity derives from expressed peptides, but the structures of the active peptides are deduced from the corresponding genomic DNA sequence. The first encoding of synthetic combinatorial libraries employed DNA as the code. A variety of other forms of encoding have been reported, including encoding with sequenceable bio-oligomers (e.g., oligonucleotides and peptides), and binary encoding with additional non-sequenceable tags.

Tagging with sequenceable bio-oligomers

The principle of using oligonucleotides to encode combinatorial synthetic libraries was described in 1992 (Brenner et al. (1992) PNAS 89:5381-5383), and an example of such a library appeared the following year (Needles et al. (1993) PNAS 90:10700-10704). A combinatorial library of nominally 7^7 (= 823,543) peptides composed of all combinations of Arg, Gln, Phe, Lys, Val, D-Val and Thr (three-letter amino acid code), each of which was encoded by a specific dinucleotide (TA, TC, CT, AT, TT, CA and AC, respectively), was prepared by a series of alternating rounds of peptide and oligonucleotide synthesis on solid support. In this work, the amine linking functionality on the bead was specifically differentiated toward peptide or oligonucleotide synthesis by simultaneously preincubating the beads with reagents that generate protected OH groups for oligonucleotide synthesis and protected NH_2 groups for peptide synthesis (here, in a ratio of 1:20). When complete, the tags each consisted of 69-mers, 14 units of which carried the code. The bead-bound library was incubated with a fluorescently labeled antibody, and beads containing bound antibody that fluoresced strongly were harvested by fluorescence-activated cell sorting (FACS). The DNA tags were amplified by PCR and sequenced, and the predicted peptides were synthesized. Following such techniques,

compound libraries can be derived for use in the subject method, where the oligonucleotide sequence of the tag identifies the sequential combinatorial reactions that a particular bead underwent, and therefore provides the identity of the compound on the bead.

The use of oligonucleotide tags permits exquisitely sensitive tag analysis. Even so, the method requires careful choice of orthogonal sets of protecting groups required for alternating co-synthesis of the tag and the library member. Furthermore, the chemical lability of the tag, particularly the phosphate and sugar anomeric linkages, may limit the choice of reagents and conditions that can be employed for the synthesis of non-oligomeric libraries. In preferred embodiments, the libraries employ linkers permitting selective detachment of the test compound library member for assay.

Peptides have also been employed as tagging molecules for combinatorial libraries. Two exemplary approaches are described in the art, both of which employ branched linkers to solid phase upon which coding and ligand strands are alternately elaborated. In the first approach (Kerr JM et al. (1993) J Am Chem Soc 115:2529-2531), orthogonality in synthesis is achieved by employing acid-labile protection for the coding strand and base-labile protection for the compound strand.

In an alternative approach (Nikolaiev et al. (1993) Pept Res 6:161-170), branched linkers are employed so that the coding unit and the test compound can both be attached to the same functional group on the resin. In one embodiment, a cleavable linker can be placed between the branch point and the bead so that cleavage releases a molecule containing both code and the compound (Ptek et al. (1991) Tetrahedron Lett 32:3891-3894). In another embodiment, the cleavable linker can be placed so that the test compound can be selectively separated from the bead, leaving the code behind. This last construct is particularly valuable because it permits screening of the test compound without potential interference of the coding groups. Examples in the art of independent cleavage and sequencing of peptide library members and their corresponding tags has confirmed that the tags can accurately predict the peptide structure.

Non-sequenceable Tagging: Binary Encoding

An alternative form of encoding the test compound library employs a set of non-sequencable electrophoric tagging molecules that are used as a binary code (Ohlmeyer et

al. (1993) PNAS 90:10922-10926). Exemplary tags are haloaromatic alkyl ethers that are detectable as their trimethylsilyl ethers at less than femtomolar levels by electron capture gas chromatography (ECGC). Variations in the length of the alkyl chain, as well as the nature and position of the aromatic halide substituents, permit the synthesis of at least 40 such tags, which in principle can encode 2^{40} (e.g., upwards of 10^{12}) different molecules. In the original report (Ohlmeyer et al., supra) the tags were bound to about 1% of the available amine groups of a peptide library via a photocleavable o-nitrobenzyl linker. This approach is convenient when preparing combinatorial libraries of peptide-like or other amine-containing molecules. A more versatile system has, however, been developed that permits encoding of essentially any combinatorial library. Here, the compound would be attached to the solid support via the photocleavable linker and the tag is attached through a catechol ether linker via carbene insertion into the bead matrix (Nestler et al. (1994) J Org Chem 59:4723-4724). This orthogonal attachment strategy permits the selective detachment of library members for assay in solution and subsequent decoding by ECGC after oxidative detachment of the tag sets.

Although several amide-linked libraries in the art employ binary encoding with the electrophoric tags attached to amine groups, attaching these tags directly to the bead matrix provides far greater versatility in the structures that can be prepared in encoded combinatorial libraries. Attached in this way, the tags and their linker are nearly as unreactive as the bead matrix itself. Two binary-encoded combinatorial libraries have been reported where the electrophoric tags are attached directly to the solid phase (Ohlmeyer et al. (1995) PNAS 92:6027-6031) and provide guidance for generating the subject compound library. Both libraries were constructed using an orthogonal attachment strategy in which the library member was linked to the solid support by a photolabile linker and the tags were attached through a linker cleavable only by vigorous oxidation. Because the library members can be repetitively partially photoeluted from the solid support, library members can be utilized in multiple assays. Successive photoelution also permits a very high throughput iterative screening strategy: first, multiple beads are placed in 96-well microtiter plates; second, compounds are partially detached and transferred to assay plates; third, a metal binding assay identifies the active

wells; fourth, the corresponding beads are rearrayed singly into new microtiter plates; fifth, single active compounds are identified; and sixth, the structures are decoded.

Examples

The invention, now being generally described, will be more readily understood by reference to the following examples, which are included merely for purposes of illustration of certain aspects and embodiments of the present invention and are not intended to limit the invention.

Experimental Results

A. Identification of Three Novel Class II human HDAC Enzymes. The NCBI database was screened with the yeast Hda1p amino acid sequence to identify human homologs. The complete ORFs of HDAC 4, 5, and 6 were constructed as described in Methods. HDAC4 consists of 1085 amino acids, with a putative catalytic region beginning at amino acid 802 (Figure 1A). HDAC5 is highly homologous to HDAC4 (51% identity, 63% similarity), with 1123 amino acids and a catalytic region beginning at amino acid 832 (Figure 1B). HDAC6 is the largest HDAC protein yet identified in humans, with 1216 amino acids. It is also unique in that it consists of an apparent internal dimer containing two catalytic domains, with the first beginning at amino acid 215 and the second at amino acid 610 (Figure 1C). The two catalytic regions in HDAC6 are highly homologous to each other (47% identity, 64% similarity) and therefore the protein may have arisen evolutionarily from an in frame gene duplication event. All the catalytic domains of these three proteins are highly conserved with respect to each other and previously identified HDAC proteins (Figure 1D). There are 15 invariant residues in this region between HDAC1, 4, 5, 6 proteins and Hda1p, and a total of 37 invariant residues within the four catalytic domains of the three new HDAC proteins. This level of sequence conservation strongly suggests that these novel proteins have deacetylase activity.

B. Differential Expression of Class II HDACs in Human Tissues. Northern blot analyses indicate differential tissue expression of the human class II HDACs (Figure 2).

HDAC4 is detectable as a 9.6 kb transcript in brain, heart and skeletal muscle tissues. The signal is very weak, however, which may have been caused by the fact that the transcript is at the upper size limit for mRNA samples in this particular blot. It is possible that HDAC4 is expressed in other tissues, but the quantity of transcript in these samples may be too low to detect. This finding is consistent with the small number of ESTs corresponding to this cDNA present in the database. HDAC5 expression partially overlaps that of HDAC4, and is observed primarily in brain, heart, skeletal muscle, and placental tissues as a 3.7 kb transcript. HDAC6, which is present as a 5 kb transcript, has the highest expression levels in heart, liver, kidney, and pancreas. The differences in tissue expression may reflect a tissue specific function of these enzymes.

C. *In Vitro* Histone Deacetylase Activity of Class II HDACs. To determine if HDAC4, 5 and 6 possess histone deacetylase activity, the recombinant proteins were assayed for enzymatic activity *in vitro*. Epitope-tagged recombinant HDACs 1, 4, 5, and 6 were expressed in Tag-Jurkat cells and immunoprecipitated. The immunoprecipitates were incubated with ^{3}H-acetate labeled histones, and the subsequent release of ^{3}H-acetate was quantified by scintillation counting. All four HDAC enzymes exhibit deacetylation activity that is at least two-fold above background levels (Figure 3A). In each case, activity is greatly reduced by the presence of 300 nM TSA, a potent HDAC inhibitor. HDAC1 and HDAC6 possess comparable activity, whereas that of HDAC4 and HDAC5 is somewhat reduced. This is most likely due to lower expression levels for these two recombinant proteins rather than inherently weaker histone deacetylase activity (see Fig 5A). Furthermore, co-immunoprecipitation experiments (see Fig 5b) demonstrate the association of HDAC4 and HDAC5 with HDAC3. It is possible that the observed HDAC activity is due to HDAC3. However, sequence analysis suggests that HDAC4 and HDAC5 possess functional catalytic domains, and therefore should contribute to the activity. Therefore, *in vitro*, all three human class II HDACs can deacetylate histones in a trichostatin-sensitive manner.

Previously, Hda1p was shown to preferentially deacetylate histone H3 *in vitro* (16). In order to determine if HDAC4, 5 and 6 display similar preferences, the immunopurified recombinant proteins were incubated with ^{3}H-acetate labeled histones

and the deacetylase reactions were separated by SDS/PAGE to identify the different histone isotypes. The gel was then exposed to autoradiography to determine the relative amount of acetylated histones remaining in each case. HDAC1, 4, 5, and 6 deacetylate all four core histones equally well, though again deacetylation by HDAC4 and HDAC5 is incomplete (Figure 3B).

D. Independently Active Catalytic Domains of HDAC6. HDAC6 possesses two separate putative catalytic domains. Site-directed mutagenesis was performed in order to determine if either domain required the other for catalytic activity. Histidine 141 of HDAC1 was shown previously to be critical for deacetylase activity (18). The corresponding histidine residues in each catalytic domain of HDAC6 were mutated to alanine, to produce the H216A and H611A single mutants and the H216/611A double mutant. The mutant HDAC6 proteins were expressed and assayed for *in vitro* deacetylation of histones. Mutation of either H216 or H611 to alanine results in a slight reduction of histone deacetylase activity, and simultaneous mutation of both sites abrogates this activity completely (Figure 4A, 4B). Furthermore, a truncation of HDAC6 containing the N-terminal 460 amino acids, and therefore only the first catalytic domain, is still catalytically active (data not shown). Therefore, both catalytic domains of HDAC6 are fully functional histone deacetylases and contribute independently to the overall activity of the wild-type HDAC6 protein.

E. Expression and Co-Immunoprecipitation of Class II HDACs. Recombinant, FLAG epitope-tagged proteins were subjected to Western blot analysis to address possible protein-protein interactions. HDAC1 migrates as a band slightly above its expected size of 55 kDa in SDS/PAGE (Figure 5A). Recombinant HDAC4 and 6 appeared above their theoretical molecular weights of 119 kDa and 131 kDa, respectively. The expression level of HDAC5 is significantly lower than the others, and the protein appears as a doublet, both in the lysate (data not shown) and in the immunoprecipitate (Fig 5A). This doublet may be the result of post-translational modifications or partial proteolytic degradation. The high molecular weight diffuse signal apparent in the blot is most likely due to cross-reaction of the secondary mouse antibody with contaminating

FLAG antibody used for the immunoprecipitation. This signal is partially masked by the comigration of the recombinant HDAC4, 5 and 6 proteins.

HDAC1 has been shown to be associated with a variety of transcription-related proteins, including the CHD chromodomain proteins, metastasis-associated factors (MTA), the co-repressor mSin3A and the histone-binding protein RbAp48. In order to determine if HDAC4, 5, and 6 associated with the same proteins *in vivo*, a series of co-immunoprecipitation experiments was performed. Immunoprecipitates were probed with -CHD4, -mSin3A, -MTA, -RbAp48, -HDAC1, and -HDAC3 antibodies (Figure 5B). The HDAC1 sample contains bands corresponding to all proteins with the exception of HDAC3, as anticipated. There is a band in the mSin3A blot of the HDAC4 immunoprecipitate, which appears at a lower molecular weight than expected for mSin3A. The nature of this band is unclear, since it does not correspond to any of the previously observed forms of mSin3A. HDAC4 coimmunoprecipitates with Rbp48 and HDAC3, though none of the other proteins were apparent. HDAC5 associates only with HDAC3, though it is possible that the expression levels were too low to detect other associated factors. HDAC6 does not appear to interact with any of these proteins, despite robust expression, nor was HDAC1 or HDAC2 found to co-immunoprecipitate with the class II HDACs. This analysis suggests that these novel class II HDAC proteins are biochemically distinct from HDAC1 *in vivo*.

F. Immunoprecipitation of HDAC4 and HDAC5 complexes

Recombinant, FLAG-epitope tagged HDAC1 and HDAC4 were transiently expressed in TAg Jurkat cells and immunoprecipitated with -FLAG antibodies. The purified proteins were separated by SDS/PAGE and visualized by silver staining. Comparison with the mock-transfected negative control and the HDAC1 samples revealed the presence of specific 30- and 32-kDa protein bands in the HDAC4 immunoprecipitate. Peptides derived from these proteins were analyzed by peptide microsequencing and found to correspond to the and isoforms of 14-3-3, respectively (Figure 7A). Due to the high degree of sequence similarity between HDAC4 and HDAC5 (51% identity), it was hypothesized that HDAC5 associates with 14-3-3 proteins as well. The presence of these two 14-3-3 protein isoforms in both HDAC4 and

HDAC5 immunoprecipitates was confirmed by Western blot analysis with isoform-specific antibodies (Figure 7B). This analysis also confirmed the previously observed association of HDAC3 with both class II HDACs. These immunoprecipitation experiments suggest that HDAC4 and HDAC5 can associate, either directly or indirectly, both with HDAC3, which is nuclear (Emiliani et al., 1998), and 14-3-3 proteins, which are generally cytoplasmic (Burbelo and Hall, 1995).

G. NUCLEAR-CYTOPLASMIC SHUTTLING OF HDAC4 AND HDAC5 IS REGULATED BY BINDING TO 14-3-3

We and other groups (Miska et al., 1999) have observed by immunofluorescence that HDAC4 and HDAC5 can be localized to either the cytoplasm or the nucleus, often aggregating in discrete foci (Figure 8A). This nuclear and cytoplasmic localization is dynamic and shuttling can occur under normal conditions of cell growth. Recombinant HDAC4-EGFP was transiently expressed in COS-7 cells, and the localization of the protein was monitored over a period of three hours in live cells. While the localization of HDAC4-EGFP remained static in the majority of cells, shuttling was observed in some cases (data not shown). This nuclear-cytoplasmic shuttling process could explain the apparent discrepancy observed in the immunoprecipitation experiments, in which HDAC4 and HDAC5 were found to interact with both nuclear HDAC3 and cytoplasmic 14-3-3.

Several cases have been reported in which proteins are sequestered in the cytoplasm by binding to 14-3-3, and disruption of this interaction allows the proteins to translocate into the nucleus and perform their function (Brunet et al., 1999; Lopez-Girona et al., 1999; Wang et al., 1999; Yang et al., 1999). It is possible that binding of HDAC4 and HDAC5 to 14-3-3 sequesters these proteins in the cytoplasm, where they are presumably unable to repress transcription. Upon loss of 14-3-3 binding, HDAC4 and HDAC5 could translocate into the nucleus, bind to HDAC3 and MEF2, and repress MEF2-dependent gene expression. In order to study the effect of 14-3-3 binding on HDAC4 localization, HDAC4-EGFP and myc-tagged 14-3-3 were co-expressed in U2OS cells, and the cellular localization of HDAC4 was analyzed by immunofluorescence in triplicate experiments (Figure 8B). Expression of HDAC4-EGFP

alone results in a cytoplasmic localization of HDAC4 in 67 (± 3) % of the cells, while simultaneous overexpression of 14-3-3 increases this to 97 (± 1)% of the cells. This correlation suggests that 14-3-3 may play a role in sequestering HDAC4 in the cytoplasm.

H. ASSOCIATION OF HDAC4 AND HDAC5 WITH 14-3-3 AND HDAC3

It is known that 14-3-3 proteins bind to phosphorylated serine or threonine residues in defined consensus sequences (Rittinger et al., 1999; Yaffe et al., 1997). It was hypothesized that phosphorylation of HDAC4 and HDAC5 would allow association with 14-3-3 and sequestration in the cytoplasm. Dephosphorylation of these HDACs should result in the loss of interaction with 14-3-3, with subsequent translocation to the nucleus and binding to HDAC3. To test this hypothesis, the effect of phosphorylation of HDAC4 and HDAC5 on their association with14-3-3 and HDAC3 was analyzed.

Recombinant FLAG-tagged HDAC4 and HDAC5 were transiently expressed in TAg Jurkat cells that were subsequently treated with the general serine/threonine kinase inhibitor staurosporine or the PP1 and PP2A phosphatase inhibitor calyculin A. The recombinant proteins were immunoprecipitated and analyzed by western blot (Figure 9A). Under dephosphorylating conditions due to staurosporine treatment, there is a decrease in binding of HDAC4 and HDAC5 to either 14-3-3 isoform, and a corresponding increase in HDAC3 association. In addition, an increase in the overall HDAC activity of the purified complex was observed (Figure 9A). Similarly, under hyper-phosphorylating conditions due to calyculin A treatment, HDAC4 and HDAC5 undergo a notable electrophoretic mobility shift, probably due to direct phosphorylation. An increase in 14-3-3 binding is observed as well, with a concomitant loss of interaction with HDAC3. This loss of binding to HDAC3 presumably causes the dramatic reduction in immunoprecipitated HDAC activity that is observed, though the activity of isolated HDAC4 and HDAC5 is still above background. These data suggest that binding of 14-3-3 to HDAC4 and HDAC5 is dependent on the presence of phosphorylated serine or threonine residues, and corresponds to a loss of interaction with HDAC3 in the nucleus.

I. Binding of 14-3-3 to HDAC4 blocks binding of importin

14-3-3 proteins have been shown to sequester *Xenopus* Cdc25 in the cytoplasm by binding near a bipartite nuclear localization sequence and blocking its interaction with the importin heterodimer (Yang et al., 1999), which is required for import into the nucleus (Gorlich, 1997). HDAC4 also contains a nuclear localization sequence (Hicks and Raikhel, 1995). In order to determine if the interaction with 14-3-3 blocks binding of the importin heterodimer, recombinant FLAG-tagged HDAC4 was immunoprecipitated from untreated, staurosporine-, and calyculin A-treated TAg-Jurkat cells and analyzed for binding to importin by Western blotting. Upon binding to 14-3-3 due to calyculin A treatment, HDAC4 fails to associate with importin (Figure 9B). Thus, sequestration of HDAC4 and HDAC5 in the cytoplasm by 14-3-3 may be caused by masking a nuclear localization sequence.

J. Mutations of the 14-3-3 binding sites cause increased nuclear localization of HDAC4

14-3-3 proteins bind to well-defined consensus sequences containing phosphorylated serine or threonine residues (Rittinger et al., 1999; Yaffe et al., 1997). There are four canonical 14-3-3 binding sites in HDAC4, three of which are well conserved in HDAC5. The serine residues in each of these three sites in HDAC4 (S246, S466, S632) were mutated to alanine in order to prohibit phosphorylation and thus prevent 14-3-3 binding. Mutation of individual sites or two sites is not sufficient to abrogate binding of 14-3-3, but mutation of all three serine residues to alanine (HDAC4 S246/466/632A) abolishes binding to 14-3-3 and , even under hyperphosphorylating conditions due to calyculin A treatment. Furthermore, localization of the triple mutant to the cytoplasm is dramatically decreased compared to the wild-type and single mutants, consistent with a role for 14-3-3 in sequestration of HDAC4 and HDAC5 in the cytoplasm.

K. Disruption of HDAC3 and HDAC4 interaction upon calyculin A treatment

Interestingly, despite its inability to bind 14-3-3 and its concomitant nuclear localization, the HDAC4 S246/466/632A triple mutant is still unable to bind to HDAC3

under calyculin A treatment. Hence, the inability of HDAC4 to associate with HDAC3 under hyperphosphorylating conditions cannot simply be due to its sequestration in the cytoplasm. Other possibilities include a change in HDAC4 conformation or electrostatic properties upon direct phosphorylation that prevents association with HDAC3, direct phosphorylation of HDAC3 resulting in similar alteration in its conformation or electrostatic nature, or modifications of additional factors that may be required for mediating the HDAC4-HDAC3 interaction.

In order to distinguish between these possibilities, immunoprecipitated HDAC4 from untreated, staurosporine-treated, or calyculin A-treated cells was incubated with lysates from untreated or calyculin A-treated TAg Jurkat cells. Briefly, forty-eight hours after transfection with HDAC4-FLAG, TAg Jurkat cells were treated with staurosporine or calyculin A for 1.5 hours. HDAC4-FLAG was immunopurified and the sample was split into thirds. One-third of the immunopurified protein was prepared for Western blot analysis, one-third was incubated for 1 hour with untreated TAg Jurkat lysate, and the remaining one-third of the sample was incubated with calyculin A-treated TAg Jurkat lysate for 1 hour. These samples were analyzed for 14-3-3 and HDAC3 binding by Western blot analysis. Untreated and staurosporine-treated (presumably hypophosphorylated) HDAC4 associated with HDAC3 under all conditions, which is consistent with previous observations. Notably, there is an increase in the amount of associated HDAC3 upon incubation with lysate from untreated cells, but not after incubation with lysate from cells treated with calyculin A. This suggests that HDAC3 from untreated cells, but not from calyculin A-treated cells, is still competent to bind HDAC4. Furthermore, there is no factor in calyculin A-treated cells that can disrupt the HDAC4-HDAC3 interaction once it has formed.

As previously observed, calyculin A-treated (hyperphosphorylated) HDAC4 does not bind to HDAC3 under normal immunoprecipitation conditions. However, upon incubation with untreated cell lysates, calyculin A-treated HDAC4 does pull down HDAC3. This interaction is not present upon incubation with calyculin A-treated lysate. Hence, calyculin A-treatment of HDAC4 does not cause it to undergo a conformational change that would prevent it from binding to HDAC3. Note that though there does seem

to be a slight decrease in the amount of HDAC3 present in the calyculin A treated lysate, it is not significant enough to explain the complete lack of binding to HDAC4.

These experiments suggest that while sequestration of HDAC4 in the cytoplasm by 14-3-3 may serve to prevent its interaction with HDAC3, this association is abrogated by an additional mechanism when the cell is exposed to hyperphosphorylating conditions. Since calyculin A-treated HDAC4 still is capable of association with HDAC3, this loss of interaction instead may be due to the phosphorylation of HDAC3 or to the modification of additional proteins required for mediating the HDAC3-HDAC4 association.

L. Increased nuclear localization of HDAC4 enhances MEF2-dependent repression

The sequestration of HDAC4 and HDAC5 in the cytoplasm by 14-3-3 presumably prevents these proteins from repressing gene transcription, and therefore represents a novel mechanism for controlling HDAC activity. In order to determine if cellular localization affects HDAC4-mediated transcriptional repression by MEF2, the following series of reporter gene assays was performed. TAg-Jurkat cells were transfected with a MEF2-luciferase reporter, the MEF2D transcription factor, and either wild-type HDAC4 or the HDAC4 S246/466/632A triple mutant, which no longer binds to 14-3-3 and displays enhanced nuclear localization. Equivalent expression levels of the wild-type and triple mutant HDAC4 recombinant proteins was confirmed by Western blot analysis (data not shown). Expression of wild-type HDAC4 decreases MEF2-dependent transcription slightly, while expression of the HDAC4 triple mutant completely represses transcription. These data are consistent with a model in which loss of association with 14-3-3 results in nuclear localization of HDAC4 and association with MEF2 and HDAC3, causing increased transcriptional repression.

M. Screening for specific inhibitors of human HDACs

Three classes of histone deacetylase enzymes have been identified in mammalian cells Gray et al. (2001) Exp Cell Res 262:75-83). Class I and Class II HDACs are structurally very similar and possess a conserved catalytic domain of approximately 390 amino acids (Finnin et al. (1999) Nature 401: 188-93). Of these, four class I (HDAC1, 2, 3 and 8) and four class II (HDAC4, 5, 6, 7) have been identified in humans. A third class

of seven human histone deacetylases with homology to yeast Sir2 has recently been characterized. These HDACs posses a very different primary sequence, and have a conserved catalytic domain of approximately 275 AA. Unlike the class I and class II HDACs, the Sir2 enzymes require NAD as a substrate for deacetylation. Furthermore, these enzymes are not affected by any known HDAC inhibitors. Thus, class I and II HDACs presumably function by the same mechanism, while the Sir2 HDACs evolved separately and function by a very different mechanism.

The class I and class II HDACs are found in several different protein complexes (rev in Ng and Bird (2000) TIBS 25: 121-6; Ahringer (2000) TIGS 16: 351-6) Kao et al (2000) Genes Dev 14: 55-66; Huang et al. (2000) Genes Dev 14: 45-54). With the exception of HDAC6, all of these HDACs are found in the nucleus (unpublished observations) and presumably function in deacetylating histones and silencing gene expression when recruited to chromosomal domains or promoters of specific genes. Several transcription factors, including nuclear hormone receptors (Xu et al. (1999) Curr Opin Genet Dev 9: 140-7) and methylated DNA-binding proteins (Ballestar and Wolffe (2001) Eur J Biochem 268: 1-6) have been shown to recruit specific HDAC complexes, but thus far it has not been possible to determine the complete set of gene targets for each of the HDACs in mammalian cells, since profiling of knockouts have not been reported. Thus, it would be highly advantageous to acquire a set of inhibitors that are specific for each of the known class I and class II HDACs. Analysis of the differences in the structures of the HDACs and previous reports of specificity of certain natural and synthetic HDAC inhibitors suggest that this is a feasible goal.

Structure of class I and II HDACs

The crystal structure of a bacterial HDAC homolog, HDLP (histone-deacetylase like protein), has been reported (Finnin et al. (1999) Nature 401: 188-93). HDLP has a single domain structure related to the open α/β class of folds. It contains a central eight-stranded parallel β-sheet, with four α-helices packed on either face. Eight additional α-helices and large loops in the β-sheet further extend the structure, and result in the formation of a deep, narrow pocket, with an adjacent internal cavity. HDLP requires Zn^{2+} for activity, and this zinc cation positioned near the bottom of the pocket, which thus presumably is the active site (see Figure 10). The active site contains several

aspartate and histidine residues, which appear to function in both coordinating the zinc ion and acting as a charge relay system to increase the nucleophilicity of a bound water molecule, thereby activating it to attack the amide bond of the acetyl-lysine (see Figure 11). The channel leading to the active site is surrounded by hydrophobic residues, which is presumably where the aliphatic chain of the acetyl-lysine residue is nestled.

A small molecule inhibitor of histone deacetylases, trichostatin A, has been co-crystallized with HDLP and allows for an analysis of the structural properties of these inhibitors. TSA contains a cap group, and aliphatic chain and a terminal hydroxamic acid functional group (Figure 12). The hydroxamic acid coordinates the zinc cation in a bidentate fashion and hydrogen bonds with some of the active site residues. The aliphatic chain fits snugly into the channel, making several van der Waals contacts with the channel residues. The cap group contacts residues on the rim of the pocket, and probably mimics the amino acids adjacent to the acetylated lysine residue in the histone. Binding of TSA causes a conformational change in a tyrosine residue on this rim (Tyr91), thereby allowing tighter packing of the cap group.

The proposed mechanism for the deacetylation reaction is similar to that seen in metallo- and serine proteases (see Figure 11). The carbonyl oxygen of the N-acetyl amide bond is thought to coordinate to the zinc cation, thereby positioning it closely to a bound water molecule and activating it for a nucleophilic attack by the water. The nucleophilicity of the water molecule, in turn, is enhanced by an interaction with the negatively charged histidine-arginine pair. Attack of the water molecule on the carbonyl carbon produces an oxy-anion intermediate, which is presumably stabilized by interacting with the zinc ion, and possibly by hydrogen bonding to a nearby tyrosine. The collapse of this intermediate would result in the break of the carbon-nitrogen bond, with the nitrogen accepting a proton from the histidine residue, and thereby producing the observed acetate and lysine products.

Structural differences of the human HDACs

The catalytic domains of the eight known human HDACs is very well conserved, but there are certain differences that may allow for the production of specific inhibitors. Most of the residues seen in the HDLP structure that interact directly with TSA are completely conserved among all of the HDACs, but the conservation in the surrounding

residues is less, with significant differences apparent between the class I and class II HDACs (Figure 13). Notably, there is significant divergence in the region of Tyr91 of HDLP, and this tyrosine residue itself is very poorly conserved among the human HDACs. This is particularly striking in that this residue is positioned on the rim of the channel and interacts directly with the cap group of TSA, and it is the only residue that shifts its conformation upon TSA binding. Thus, the considerable diversity in the region of the protein interacting with the cap group of the inhibitors suggests that it will be possible to develop specific inhibitors by altering the structure of this cap group.

Synthetic HDAC inhibitors

Several non-natural inhibitors of histone deacetylases have been synthesized, and analysis of these will facilitate the rational design of novel and specific inhibitors. The basic structure of these small molecules mimics that of TSA, in that they possess a cap group, an aliphatic chain, and a functional group that would interact with the active site. It appears that the optimal length for the aliphatic chain is five to six carbon residues, with inhibitory activity decreasing rapidly for longer and shorter chains (Jung et al. (1999) J Med Chem 42: 4669-79) There are several possible functional groups, including epoxides, which appear to form irreversible covalent bonds with active site residues (see Furumai et al. (2001) Proc Natl Acad Sci USA 98: 87-92) as well as hydroxamic acid (Jung et al (1999) J Med Chem 42: 4669-79; Furumai (2001) Proc Natl Acad Sci USA 98: 87-92; Richon et al. (1996) Proc Natl Acad Sci USA 93: 57605-8; and Richon et al. (1998) 95: 3003-7) and *o*-aminoanilines (Suzuki et al. (1999) J Med Chem 42: 3001-3) which presumably coordinate the zinc cation. Two types of cap groups are found in natural HDAC inhibitors, a small flat or linear group or a cyclic tetrapeptide. Interestingly, the type of cap group alters the affinity with which the inhibitor binds to the active site. Molecules with a small cap are not active if the functional group is a carboxylic acid, while molecules with cyclic tetrapeptide cap groups and a carboxylic acid functional group are potent inhibitors.

Generation of Specificity

Alterations in the structure of the cap group has led to specific inhibitors of HDACs in two cases. First, N-alkylations of the indole residue in the cap group of apicidin has led to the development of apicidin derivatives that are ~20-fold more potent

inhibitors of malarial HDACs than human HDAC1 (Meinke et al. (2000) J Med Chem 43: 4919-22). Secondly, HDAC6 is not inhibited by any of the cyclic tetrapeptide inhibitors, even when a hydroxamic acid is used as a functional group. Notably, HDAC6 has the greatest divergence from the other HDACs in the rim region surrounding Tyr91 in HDLP.

Design of Screen

Combinatorial libraries will be screened for specific inhibitors of HDACs 1, 4 and 6. HDAC1 was chosen to be the representative of the class I HDACs, while HDAC4 is the representative for the class II HDACs. HDAC6 will be screened because it is structurally divergent from the other class II HDACs, both in primary sequence in the catalytic domain, and the fact that it has two catalytic domains. Furthermore, HDAC6 is clearly divergent from the other HDACs in that it is not inhibited by the cyclic tetrapeptide class of inhibitors.

Thus far, an unbiased library (Sternson et al. (2001) J Am Chem Soc 123: 1740-7) has been screened for specific inhibitors of these three HDACs. In the future, libraries structurally biased to be similar to known HDAC inhibitors will be generated. In these cases, an aliphatic chain with a terminal function group, either a hydroxamic acid, carboxylic acid or an *o*-aminoaniline, will be attached to a structurally diverse core molecule. Thus, the diversity will be localized primarily to the cap group of these potential inhibitors, which is desirable in that changes in this region have been shown to lead to specificity.

Recombinant, purified FLAG-epitope tagged HDAC1 and HDAC6, as well as the catalytic domain of HDAC4(632-1070) fused to GST were incorporated in baculovirus vectors and expressed in Sf9 cells. Expression by baculovirus rather than expression in mammalian cells allowed for the isolation of larger quantities of protein, and also permitted the isolation of these HDACs without associated mammalian proteins. These proteins were labeled with Cy5 dye and used to screen a 1200 member library that had been printed on glass slides (MacBeath et al. (1999) J Am chem Soc 96: 4868-73). The molecules that were identified as reproducibly binding to any of the three HDACs were then tested in a secondary HDAC activity assay (Taunton et al. (1996) Science 272: 408-

11) None of the identified hits were able to significantly inhibit HDAC activity at 100 μM.

Material and Methods

Construction of Baculoviruses

cDNA encoding HDAC1 with a C-terminal FLAG epitope tag was cloned into the transfer vector pVL1392 (Pharmingen) and used to produce recombinant baculovirus as described previously (Hassig et al. (1997) Cell 89: 341-7). cDNA encoding HDAC6 with a C-terminal FLAG epitope tag was cloned into the NotI/XbaI site of pCDNA6-V5-HisA vector (Invitrogen). This construct, with the C-terminal FLAG, V5, and HisA tags was subcloned into the NotI/PmeI sites of pVL1392. This construct was used to generate recombinant baculovirus using the Baculogold DNA, according to the manufacturer's instructions (Pharmingen). HDAC4 recombinant baculovirus was constructed in the pAcSG2T vector, and was obtained from the Pavletich group (Memorial Sloan-Kettering Cancer Research Center, New York City). cDNA encoding residues 632 to 1070 of HDAC4 was inserted into the BamHI and EcoRI sites, in frame with the N-terminal GST coding sequence, separated by a thrombin cut site.

Expression of protein in Sf9 cells

Baculovirus was amplified to a concentration of ~$1x10^6$ pfu/ml. Sf9 cells were infected at a concentration of $1x10^6$ cells/ml, with a multiplicity of infection of 2, at the National Cell Culture Center (Minneapolis, MN). 72 hours post-infection, cells infected with HDAC1 and HDAC6 baculovirus were pelleted, washed twice in PBS and flash frozen. In the case of the HDAC4 baculovirus, the protein is secreted into the media, and thus the cells were pelleted and the media saved and stored at 4°C.

Purification of protein

In the case of GST-HDAC4(632-1070), the protein is secreted into the media. 500 μl of glutathione-agarose beads were washed in PBS and incubated with 500 mL of media, with shaking for 1.5 hours. The beads were pelleted (1000 rpm, 5') and washed twice with cold PBS. The protein was eluted with 10 mM glutathione in JLB (50 mM Tris-HCL, pH 8/150 mM NaCl/10% glycerol/0.5% TritonX-100). The beads were incubated twice with 1 mL of glutathione/JLB for 20 minutes, and once with 0.5 mL for 10 minutes. The collected eluant was dialyzed to remove the glutathione in a 10 kD

MWCO Slide-A-Lyzer (Pierce) overnight against 1 L of PBS/10% glycerol/ 2.5mM DTT. Note it was necessary to use a buffer without free amines in order to be compatible with the subsequent Cy5-labeling procedure.

Pellets of 250 and 100 mL of Sf9 cells infected with HDAC1 and HDAC6 baculovirus were lysed in 10 mL of JLB, with incubation at 4C for 20 minutes. The lysate was clarified by pelleting the cellular debris at 14,000 rpm for 15 minutes. The protein was immunoprecipitated by incubating with 100 μl of anti-FLAG M2 agarose (Sigma). The beads were washed three times for three minutes in cold MSWB (50 mM Tris-HCL, pH 8/150 mM NaCl/1 mM EDTA/0.1% Nonidet P-40) and the protein was eluted by incubating the beads in 500 μL JLB with 1 mg of FLAG peptide for four hours. The eluant was collected and the beads washed with 200 μl JLB, which was added to the eluant. The protein was dialyzed overnight against JLB in a 10 kD MWCO Slide-A-Lyzer (Pierce). Since this buffer contains amines, it was necessary to exchange the JLB with PBS/10% glycerol/ 2.5mM DTT, using a 10 kD MWCO Microcon (Millipore, Bedford, MA), according to the manufacturer's instructions.

Expression and immunoprecipitation of HDACs from mammalian cells

Constructs containing HDAC1 and HDAC6 with C-terminal FLAG epitope tags in the pBJ5 mammalian expression vector have been described (Grozinger et al. (1999) 96: 4863-73). These were transfected into TAg Jurkat cells by electroporation, and forty-eight hours later cells were lysed in JLB (50 mM Tris HCl pH 8/150 mM NaCl/10% glycerol/0.5% TritonX-100) containing a complete protease inhibitor cocktail (Boehringer-Mannheim). Lysis proceeded for 10 minutes at 4°C, after which the cellular debris was pelleted by centrifugation at 14K for 5 minutes. Recombinant proteins were immunoprecipitated from the supernatant by incubation with α-FLAG M2 agarose affinity gel (Sigma) for 1 hours at 4°C. For enzyme activity assays, the beads were washed three times with JLB at 4°C.

HDAC Assays.

^{3}H-acetate-incorporated histones were isolated from butyrate-treated HeLa cells by hydroxyapitite chromatography as described(Tong et al. (1998) Nature 395: 917-21). Immunoprecipitates were incubated with 1.4 μg (10,000 dpm) histones for three hours at

37°C. HDAC activity was determined by scintillation counting of the ethyl-acetate soluble ^3H acetic acid (Taunton et al. (1996) Science 272: 408-11).

Cy5-Labeling of Proteins

1 mg of Cy5 monofunctional reactive dye (Amersham) was resuspended in 1 mL of 50 mM $NaCO_3$, pH 8.5. Protein was added to this to a concentration of ~1 mg/mL, and the reaction was incubated for 1.5 hours at 4°C. The solution was dialyzed in a 10 kD MWCO Slide-A-Lyzer (Pierce) against 1 L of PBS/10% glycerol/2.5 mM DTT overnight, and then against 2 L of PBS/10% glycerol/2.5 mM DTT/1 mM EDTA for five hours.

Screening of slides

The 1,3 dioxane library (Sternson et al. (2001) J Am Chem Soc 123: 1740-7) was printed on glass slides as described previously. The slides were blocked for five hours in 3% BSA in PBST (0.1% Tween-20/PBS) and washed in PBST. Solutions of 400 nM of Cy5-labeled in 1% BSA/PBST were then added and incubated at 4C for 1 hour. The slides were washed three times with PBST, rinsed with ddH_2O and spun dry (800 rpm, 30 seconds). The slides were scanned on a Applied Precision ArrayWorx scanner.

Results

Expression and activity of recombinant HDACs

HDAC1-F, GST-HDAC4(632-1070), and HDAC6-F-V5-HisA were expressed in Sf9 cells and purified by FLAG-agarose beads or glutathione agarose. These proteins were then labeled with Cy5 dye and purified from the free dye by dialysis. The purified proteins were subjected to SDS-PAGE and visualized by silver stain or western blotting with α-FLAG antibody. All of the proteins expressed well, and there were no contaminants at stoichiometric concentrations apparent from the silver stain, and thus these are anticipated to be relatively pure. Note that expression in Sf9 cells allows for the purification of these proteins in the absence of the associated mammalian proteins. This is critical for the identification of specific HDAC inhibitors, since HDAC1 usually associates with HDAC2 in most mammalian cell lines, while HDAC4 associates with HDAC3.

Screening on printed 1,3 dioxane library with recombinant HDACs

Approximately 2000 compounds from the 1,3 dioxane library were printed on slides (MacBeath et al. (1999) J Am Chem Soc 96: 4868-73) and screened with the three Cy5-labeled HDACs in triplicate. Compounds that reproducibly bound these HDACs were identified and their location in the arrayed plates was determined. An example of the observed results for an HDAC6-Cy5 screen (compound 4-P9) is shown in Figure 15. The following compounds were identified in this assay:

HDAC1/HDAC6 common hits: 2-N20, 4-J18, 7-B2, 7-M1, 11-A15

HDAC1 unique hits: 4-H12

HDAC6 unique hits: 2-M7, 4-M9, 4-P9

Retest hits on HDAC1 and HDAC6

FLAG-tagged HDAC1 and HDAC6 were expressed in TAg Jurkat cells and purified by immunoprecipitation using the anti-FLAG antibody. Note that TAg Jurkat cells have undetectable quantities of soluble HDAC2, and therefore HDAC1 that is purified from these cells does not contain associated HDAC2. The proteins were used in *in vitro* HDAC assays with the compounds identified in the screen, using concentration of approximately 50 μM.

HDAC6 appeared to be inhibited by compound 11-A15, while HDAC1 was unaffected by it (see Figure 16). Thus this compound was resynthesized and retested at 100 μM concentration, in triplicate (see results, Figure 17). It did not inhibit either HDAC at this concentration. Retesting of other compounds also demonstrated that the previously observed inhibition was not reproducible. This suggests that the original results were due to the low signal-to-noise ratio.

Conclusions

In this screen, three different HDACs were chosen to represent the eight known human HDAC proteins. Based on sequence similarities, HDAC1 is highly related to HDACs 2 and 3, while HDAC4 is closely related to HDAC5 and 7. HDAC6 ~~is~~ has unique features, both in its sequence and in the fact that it is not inhibited by the cyclic tetrapeptide class of HDAC inhibitors.

The methodology outlined above will allow for rapidly screening several thousand printed compounds for specific binding to any of these three proteins. Positives identified in this binding assay can easily be tested in in vitro HDAC assays with the

purified proteins to screen for specific inhibitory activity. Furthermore, this activity assay can be performed in very small volumes, and thus resynthesis of the positive compounds will not be required for this secondary screen.

None of the compounds identified in the binding assay tested positively in the subsequent activity assay. This suggests that these compounds were either false positives or did not bind to the active site of the enzymes. This is not particularly unexpected, since the library that was tested was not structurally biased to mimic HDAC inhibitors. Future libraries will include moieties containing an aliphatic side chain terminating with a hydroxamic acid, carboxylic acid, or o-amino anilide group. Sample compounds from these libraries have been shown to inhibit HDAC1 activity with an IC_{50} of 5-100 μM, and thus these will probably be more suitable for these screens.

Experimental Methods

A. Cloning of HDAC4, HDAC5 and HDAC6. The amino acid sequence of yeast Hda1p was used in a tblastn search of the NCBI databases to identify human homologs of Hda1p. A cDNA clone for HDAC4 was obtained from the Kazusa DNA Research Institute, Kisarazu, Japan (Gene name KIAA0288, GenBank Accession Number 3024889). The full length clone was 8459 bp, with a predicted ORF of 2903 bp. A comparison of the HDAC4 clone with the HDAC5 sequence, however, revealed that the putative 5'UTR of HDAC4 was highly homologous to the N-terminal coding sequence of HDAC5. A truncation in the HDAC4 ORF seems to have been caused by a C to T point mutation in the putative 5'UTR of the HDAC4 clone, resulting in the conversion of a glutamate codon to a stop codon. This hypothesis was confirmed by obtaining the remaining N-terminal 352 bp by PCR from a 5'Stretch cDNA Leukemia Library (Clontech) and sequencing the product. A C-terminal FLAG epitope tag was incorporated into the complete HDAC4 ORF of 3255 bp, which was then inserted into the Not I - Eco RI sites of a mammalian expression vector (pBJ5).

An EST (GenBank Accession #R64669) homologous to yHda1p was identified and obtained from the I.M.A.G.E. Consortium (Lawrence Livermore National Laboratory). This sequence was used to generate a random primed probe (Boehringer

Mannheim) to screen a Lambda ZAP II Jurkat cDNA library (Stratagene). A 2.3 kb cDNA clone was identified that contained a partial ORF of HDAC5. A blastn search of the NCBI database facilitated the assembly of the full-length cDNA sequence as a combination of a second clone from the Kazusa DNA Research Institute (Gene Name KIAA0600, GenBank Accession #3043724) and the cDNA clones containing the 11-js mRNA sequence (GenBank Accession # AF039241), kindly provided by Jeff Swensen, (University of Utah). An ORF of 3369 bp was identified and assembled into a C-terminal FLAG construct by subcloning into the Not I – Xho I sites of a pBJ5 vector.

A second EST homologous to yHDA1p was identified and used to screen the Lambda ZAP II Jurkat cDNA library. The sequence of the 2.5 kb clone produced was used in a blastn search to reveal the full length sequence of HDAC6 in the jm-21 mRNA sequence (GenBank Accession #AJ011972). This information was used to obtain the remaining 1.5 kb by nested PCR from a U957 cDNA library, kindly provided by Don Ayer (University of Utah) to produce the full length ORF of 3648bp. A C-terminal FLAG epitope tag was incorporated into this clone, which was inserted into the Not I – Spe I sites of pBJ5.

B. Northern blot analysis. Multiple human tissue Northern blots were obtained from Clontech. Probes were generated and blots were stripped using Strip-EZ DNA probe synthesis and removal kit (Ambion). Prehybridization and hybridization was carried out according to manufacturer's instructions using ExpressHyb solution (Clontech). For HDAC4 expression analysis, a 12-lane tissue blot was probed with the 895 bp SalI-SacI fragment of the HDAC4 gene. For HDAC5 expression analysis, an 8-lane tissue blot was probed with the 993 bp SacI-SacII fragment of the HDAC5 gene. This blot was stripped and reprobed for HDAC6 expression using the 667 bp SphI-AvrII fragment of the HDAC6 gene. Blots were stripped and reprobed with -actin cDNA as a control (Clontech).

C. Antibodies, immunoprecipitation and Western blotting. Antibodies against HDAC1 (11), HDAC3 (17) and RbAp48 (17) have been described previously. Antibodies against mSin3A were kindly provided by Don Ayer (18), antibodies to the

PHD domain of CHD4 were provided by Weidong Wang (National Institute on Aging/NIH) (6) and antibodies to the N-terminal domain of MTA were provided by Yasushi Toh (Kyushi University, Fukuoka, Japan) (19, 20).

Forty-eight hours after transfection, cells were lysed in JLB (50 mM Tris HCl pH 8/150 mM NaCl/10% glycerol/0.5% TritonX-100) containing a complete protease inhibitor cocktail (Boehringer-Mannheim). Lysis proceeded for 10 minutes at 4°C, after which the cellular debris was pelleted by centrifugation at 14K for 5 minutes. Recombinant proteins were immunoprecipitated from the supernatant by incubation with -FLAG M2 agarose affinity gel (Sigma) for 2 hours at 4°C. For Western blot analysis, the beads were washed three times for 5 minutes at room temperature with MSWB (50 mM Tris HCl pH 8/150 mM NaCl/1 mM EDTA/0.1% NP-40) and the proteins were separated by SDS/PAGE. For enzyme activity assays, the beads were washed three times with JLB at 4°C.

D. HDAC Assays. ^{3}H-acetate-incorporated histones were isolated from butyrate-treated HeLa cells by hydroxyapitite chromatography as described (4). Immunoprecipitates were incubated with 1.4 μg (10,000 dpm) histones for three hours at 37°C. HDAC activity was determined by scintillation counting of the ethyl-acetate soluble ^{3}H acetic acid (11). Inhibition of enzyme activity by trichostatin (TSA) was performed by incubating samples with 300 nM TSA (Wako) for 10 minutes prior to addition of the labeled histones.

E. Histone isolation and substrate specificity determination. For deacetylase assays, 6 g of histones were incubated with immunoprecipitated recombinant enzyme or negative control (RbAp48 transfected) for 3 hours at 37° C in HD buffer (50mM Tris, pH 8.0/150mM NaCl/10% glycerol). Reactions were stopped with SDS loading buffer and proteins were separated by 20% SDS/PAGE.

F. HDAC6 Mutagenesis. The H216A and H611A mutations were produced by PCR overlap extension. For each mutant, internal primers to both strands with the corresponding CAC to GCC mismatches were synthesized and used to amplify two overlapping fragments containing the mutation in the common region. These fragments

were then used as templates in a second PCR reaction with the same flanking external primers. The H216A fragment was ligated into the Not I and Dra III sites of the wild-type HDAC6-pBJ5 construct. The H611A fragment was ligated into the Dra III and BstE II sites of the wild-type HDAC6-pBJ5. The H216/611A double mutant was constructed by ligating the H216A fragment into the Not I and Dra III sites of the HDAC6-H611A-pBJ5 clone. Plasmids were sequenced to ensure the incorporation of the mutations.

G. Cell culture and transfections. TAg-Jurkat cells were transfected by electroporation with 5 μg of FLAG-epitope tagged pBJ5 constructs for expression of recombinant proteins. Cells were mock-transfected without DNA or with an untagged RbAp48 construct in pBJ5 as a negative control. Cells were harvested 48 hours post-transfection.

H. DNA constructs

FLAG-epitope-tagged HDAC4 and HDAC5 constructs in the pBJ5 mammalian expression vector have been described previously (Grozinger et al, 1999). The HDAC4-EGFP clone in pBJ5 was made by ligation of a *Not I –Xba I* HDAC4-FLAG fragment to a *Xba I – Sal I* EGFP fragment, which was generated by PCR from a plasmid containing EGFP (Clontech). The HDAC5-EGFP/pBJ5 construct was made similarly using *Not I – Xba I* HDAC5-FLAG fragment. The C-terminally-myc-epitope-tagged 14-3-3 was produced by PCR from a Jurkat cDNA library (Stratagene) and subcloned into the *Not I* and *Spe I* sites of pBJ5.

The S246A, S466A and S632A single, double and triple mutations in HDAC4 were generated by PCR overlap extension. For each mutant, internal primers to each complementary strand with the corresponding base pair mismatches were synthesized and used to amplify two overlapping fragments containing the mutations in the common region. These fragments were then used in a second PCR with the same flanking external primers. The sequence of the primers used will be made available on request. The fragments were cloned into the *Not I* and *Sac II* sites of the HDAC4-F/pBJ5 construct and sequenced to ensure proper incorporation of the mutations.

The reporter used in the luciferase assays contains three copies of the desmin MEF2 site in pGL2-E1b-Luciferase and was generously provided by Eric Olson

(University of Texas, Southwestern Medical Center). The myc-epitope-tagged MEF2D/ pSCT mammalian expression plasmid was kindly provided by Jun Liu (MIT).

I. Antibodies, immunoprecipitation and western blotting

Antibodies against HDAC3 have been described previously (Hassig et al., 1998). Isoform-specific antibodies against 14-3-3 and were obtained from Santa Cruz Biotechnology, while antibodies to importin (-Rch1) were acquired from Transducin Laboratories. -FLAG M2 antibodies and -mouse IgG Texas-red conjugated secondary antibodies for immunofluorescence were obtained from Sigma, while -c-myc antibodies were purchased from Upstate Biotechnology.

Forty-eight hours after transfection, cells were lysed in JLB (50 mM Tris HCl pH 8/150 mM NaCl/10% glycerol/0.5% TritonX-100) containing a complete protease inhibitor cocktail (Boehringer-Mannheim) and phosphatase inhibitors (20 mM $NaH_2(PO_4)$, pH 7.2; 25 mM NaF, 2 mM EDTA). Lysis proceeded for 10 minutes at 4°C, after which the cellular debris was pelleted by centrifugation at 14K for 5 minutes. Recombinant proteins were immunoprecipitated from the supernatant by incubation with -FLAG M2 agarose affinity gel (Sigma) for 1 hour at 4°C. For Western blot analysis and silver staining, the beads were washed three times for 5 minutes at room temperature with MSWB (50 mM Tris HCl pH 8/150 mM NaCl/1 mM EDTA/0.1% NP-40) and the proteins were separated by SDS/PAGE. For enzyme activity assays, the beads were washed three times with JLB at 4°C.

J. Peptide Microsequencing

The sequence analysis was performed at the Harvard Microchemistry Facility by microcapillary reverse-phase HPLC nano-electrospray tandem mass spectrometry (LC/MS/MS) on a Finnigan LCQ quadropole ion trap mass spectrometer,

K. HDAC Assays.

^{3}[H]-acetate-incorporated histones were isolated from butyrate-treated HeLa cells by hydroxyapitite chromatography as described (Tong et al., 1998). Immunoprecipitates were incubated with 1.4 μg (10,000 dpm) histones for two hours at 37°C. HDAC activity

was determined by scintillation counting of the ethyl-acetate soluble ^{3}[H] acetic acid (Taunton et al., 1996).

L. Cell culture and transfections.

TAg-Jurkat cells were transfected by electroporation with 5 μg of DNA for expression of recombinant proteins, or mock-transfected without DNA as a negative control. Cells were harvested 48 hours post-transfection. When required, cells were treated with 200 M staurosporin (Calbiochem) or 20 nM calyculin A (Calbiochem) for 1.5 hours prior to harvesting.

M. Immunofluorescence

Cos-7 or U20S cells were plated on coverslips and allowed to attach overnight. Subsequently they were transfected with 1-2 g of DNA using the Lipofectamine PLUS system (Gibco). Forty-eight hours later, cells were fixed with paraformaldehye and stained with antibodies and Hoechst dye (Molecular Probes), or live cells were stained with Hoechst and the EGFP was visualized directly using an fluorescence microscope (Spencer Scientific Corporation). For the time course studies, a Delta Vision confocal microscope (Applied Precision Technologies) was used.

N. Reporter gene assays

For each sample, 10 million TAg Jurkat cells were transfected with a total of 5 g of DNA. A constitutive -galactosidase expression vector was used as a control for protein expression levels in the luciferase assays. Thirty-eight hours after transfection, the samples were harvested and split into sets of three. Luciferase activity was determined according manufacturer's instructions (Promega), and -galactosidase activity was determined using a standard -galactosidase assay. Luciferase values (relative light units) were normalized for transfection efficiency by dividing by -gal activity. These assays were performed four times with similar results.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, numerous equivalents to the specific polypeptides, nucleic acids, methods, assays and reagents described herein. Such equivalents are considered to be within the scope of this invention and are covered by the following claims.

Sequence Listings

SEQ. ID. NO. 1 (HDAC4)

```
1    ATGAGCTCCC AAAGCCATCC AGATGGACTT TCTGGCCGAG ACCAGCCAGT GGAGCTGCTG
61   AATCCTGCCC GCGTGAACCA CATGCCCAGC ACGGTGGATG TGGCCACGGC GCTGCCTCTG
121  CAAGTGGCCC CCTCGGCAGT GCCCATGGAC CTGCGCCTGG ACCACCAGTT CTCACTGCCT
181  GTGGCAGAGC CGGCCCTGCG GGAGCAGCAG CTGCAGCAGG AGCTCCTGGC GCTCAAGCAG
241  AAGCAGCAGA TCCAGAGGCA GATCCTCATC GCTGAGTTCC AGAGGCAGCA CGAGCAGCTC
301  TCCCGGCAGC ACGAGGCGCA GCTCCACGAG CACATCAAGC AACAACAGGA GATGCTGGCC
361  ATGAAGCACC AGCAGGAGCT GCTGGAACAC CAGCGGAAGC TGGAGAGGCA CCGCCAGGAG
421  CAGGAGCTGG AGAAGCAGCA CCGGGAGCAG AAGCTGCAGC AGCTCAAGAA CAAGGAGAAG
481  GGCAAAGAGA GTGCCGTGGC CAGCACAGAA GTGAAGATGA AGTTACAAGA ATTTGTCCTC
541  AATAAAAAGA AGGCGCTGGC CCACCGGAAT CTGAACCACT GCATTTCCAG CGACCCTCGC
601  TACTGGTACG GGAAAACGCA GCACAGTTCC CTTGACCAGA GTTCTCCACC CCAGAGCGGA
661  GTGTCGACCT CCTATAACCA CCCGGTCCTG GGAATGTACG ACGCCAAAGA TGACTTCCCT
721  CTTAGGAAAA CAGCTTCTGA ACCGAATCTG AAATTACGGT CCAGGCTAAA GCAGAAAGTG
781  GCCGAAAGAC GGAGCAGCCC CCTGTTACGC AGGAAAGACG GGCCAGTGGT CACTGCTCTA
841  AAAAAGCGTC CGTTGGATGT CACAGACTCC GCGTGCAGCA GCGCCCCAGG CTCCGGACCC
901  AGCTCACCCA ACAACAGCTC CGGGAGCGTC AGCGCGGAGA ACGGTATCGC GCCCGCCGTC
961  CCCAGCATCC CGGCGGAGAC GAGTTTGGCG CACAGACTTG TGGCACGAGA AGGCTCGGCC
1021 GCTCCACTTC CCCTCTACAC ATCGCCATCC TTGCCCAACA TCACGCTGGG CCTGCCTGCC
1081 ACCGGCCCCT CTGCGGGCAC GGCGGGCCAG CAGGACACCG AGAGACTCAC CCTTCCCGCC
1141 CTCCAGCAGA GGCTCTCCCT TTTCCCCGGC ACCCACCTCA CTCCCTACCT GAGCACCTCG
1201 CCCTTGGAGC GGGACGGAGG GGCAGCGCAC AGCCCTCTTC TGCAGCACAT GGTCTTACTG
1261 GAGCAGCCAC CGGCACAAGC ACCCCTCGTC ACAGGCCTGG GAGCACTGCC CCTCCACGCA
1321 CAGTCCTTGG TTGGTGCAGA CCGGGTGTCC CCCTCCATCC ACAAGCTGCG GCAGCACCGC
1381 CCACTGGGGC GGACCCAGTC GGCCCCGCTG CCCCAGAACG CCCAGGCTCT GCAGCACCTG
1441 GTCATCCAGC AGCAGCATCA GCAGTTTCTG GAGAAACACA AGCAGCAGTT CCAGCAGCAG
```

1501 CAACTGCAGA TGAACAAGAT CATCCCCAAG CCAAGCGAGC CAGCCCGGCA GCCGGAGAGC

1561 CACCCGGAGG AGACGGAGGA GGAGCTCCGT GAGCACCAGG CTCTGCTGGA CGAGCCCTAC

1621 CTGGACCGGC TGCCGGGGCA GAAGGAGGCG CACGCACAGG CCGGCGTGCA GGTGAAGCAG

1681 GAGCCCATTG AGAGCGATGA GGAAGAGGCA GAGCCCCCAC GGGAGGTGGA GCCGGGCCAG

1741 CGCCAGCCCA GTGAGCAGGA GCTGCTCTTC AGACAGCAAG CCCTCCTGCT GGAGCAGCAG

1801 CGGATCCACC AGCTGAGGAA CTACCAGGCG TCCATGGAGG CCGCCGGCAT CCCCGTGTCC

1861 TTCGGCGGCC ACAGGCCTCT GTCCCGGGCG CAGTCCTCAC CCGCGTCTGC CACCTTCCCC

1921 GTGTCTGTGC AGGAGCCCCC CACCAAGCCG AGGTTCACGA CAGGCCTCGT GTATGACACG

1981 CTGATGCTGA AGCACCAGTG CACCTGCGGG AGTAGCAGCA GCCACCCCGA GCACGCCGGG

2041 AGGATCCAGA GCATCTGGTC CCGCCTGCAG GAGACGGGCC TCCGGGGCAA ATGCGAGTGC

2101 ATCCGCGGAC GCAAGGCCAC CCTGGAGGAG CTACAGACGG TGCACTCGGA AGCCCACACC

2161 CTCCTGTATG GCACGAACCC CCTCAACCGG CAGAAACTGG ACAGTAAGAA ACTTCTAGGC

2221 TCGCTCGCCT CCGTGTTCGT CCGGCTCCCT TGCGGTGGTG TTGGGGTGGA CAGTGACACC

2281 ATATGGAACG AGGTGCACTC GGCGGGGGCA GCCCGCCTGG CTGTGGGCTG CGTGGTAGAG

2341 CTGGTCTTCA AGGTGGCCAC AGGGGAGCTG AAGAATGGCT TTGCTGTGGT CCGCCCCCCT

2401 GGACACCATG CGGAGGAGAG CACGCCCATG GGCTTTTGCT ACTTCAACTC CGTGGCCGTG

2461 GCAGCCAAGC TTCTGCAGCA GAGGTTGAGC GTGAGCAAGA TCCTCATCGT GGACTGGGAC

2521 GTGCACCATG GAAACGGGAC CCAGCAGGCT TTCTACAGCG ACCCTAGCGT CCTGTACATG

2581 TCCCTCCACC GCTACGACGA TGGGAACTTC TTCCCAGGCA GCGGGGCTCC TGATGAGGTG

2641 GGCACAGGGC CCGGCGTGGG TTTCAACGTC AACATGGCTT TCACCGGCGG CCTGGACCCC

2701 CCCATGGGAG ACGCTGAGTA CTTGGCGGCC TTCAGAACGG TGGTCATGCC GATCGCCAGC

2761 GAGTTTGCCC CGGATGTGGT GCTGGTGTCA TCAGGCTTCG ATGCCGTGGA GGGCCACCCC

2821 ACCCCTCTTG GGGGCTACAA CCTCTCCGCC AGATGCTTCG GGTACCTGAC GAAGCAGCTG

2881 ATGGGCCTGG CTGGCGGCCG GATTGTCCTG GCCCTCGAGG GAGGCCACGA CCTGACCGCC

2941 ATTTGCGACG CCTCGGAAGC ATGTGTTTCT GCCTTGCTGG GAAACGAGCT TGATCCTCTC

3001 CCAGAAAAGG TTTTACAGCA AAGACCCAAT GCAAACGCTG TCCGTTCCAT GGAGAAAGTC

3061 ATGGAGATCC ACAGCAAGTA CTGGCGCTGC CTGCAGCGCA CAACCTCCAC AGCGGGGCGT

3121 TCTCTGATCG AGGCTCAGAC TTGCGAGAAC GAAGAAGCCG AGACGGTCAC CGCCATGGCC

3181 TCGCTGTCCG TGGGCGTGAA GCCCGCCGAA AAGAGACCAG ATGAGGAGCC CATGGAAGAG

3241 GAGCCGCCCC TGTAG

SEQ. ID. NO. 2 (HDAC4)

MSSQSHPDGLSGRDQPVELLNPARVNHMPSTVDVATALPLQVAPSAVPMDLRLDHQFSLPVAEPA
LREQQLQQELLALKQKQQIQRQILIAEFQRQHEQLSRQHEAQLHEHIKQQQEMLAMKHQQELLEH
QRKLERHRQEQELEKQHREQKLQQLKNKEKGKESAVASTEVKMKLQEFVLNKKKALAHRNLNH
CISSDPRYWYGKTQHSSLDQSSPPQSGVSTSYNHPVLGMYDAKDDFPLRKTASEPNLKLRSRLKQ
KVAERRSSPLLRRKDGPVVTALKKRPLDVTDSACSSAPGSGPSSPNNSSGSVSAENGIAPAVPSIPA
ETSLAHRLVAREGSAAPLPLYTSPSLPNITLGLPATGPSAGTAGQQDTERLTLPALQQRLSLFPGTH
LTPYLSTSPLERDGGAAHSPLLQHMVLLEQPPAQAPLVTGLGALPLHAQSLVGADRVSPSIHKLRQ
HRPLGRTQSAPLPQNAQALQHLVIQQQHQQFLEKHKQQFQQQQLQMNKIIPKPSEPARQPESHPEE
TEEELREHQALLDEPYLDRLPGQKEAHAQAGVQVKQEPIESDEEEAEPPREVEPGQRQPSEQELLF
RQQALLLEQQRIHQLRNYQASMEAAGIPVSFGGHRPLSRAQSSPASATFPVSVQEPPTKPRFTTGLV
YDTLMLKHQCTCGSSSSHPEHAGRIQSIWSRLQETGLRGKCECIRGRKATLEELQTVHSEAHTLLY
GTNPLNRQKLDSKKLLGSLASVFVRLPCGGVGVDSDTIWNEVHSAGAARLAVGCVVELVFKVAT
GELKNGFAVVRPPGHHAEESTPMGFCYFNSVAVAAKLLQQRLSVSKILIVDWDVHHGNGTQQAF
YSDPSVLYMSLHRYDDGNFFPGSGAPDEVGTGPGVGFNVNMAFTGGLDPPMGDAEYLAAFRTVV
MPIASEFAPDVVLVSSGFDAVEGHPTPLGGYNLSARCFGYLTKQLMGLAGGRIVLALEGGHDLTAI
CDASEACVSALLGNELDPLPEKVLQQRPNANAVRSMEKVMEIHSKYWRCLQRTTSTAGRSLIEAQ
TCENEEAETVTAMASLSVGVKPAEKRPDEEPMEEEPPL

SEQ. ID. NO. 3 (HDAC5)

1 ATGAACTCTC CCAACGAGTC GGATGGGATG TCAGGTCGGG AACCATCCTT GGAAATCCTG
61 CCGCGGACTT CTCTGCACAG CATCCCTGTG ACAGTGGAGG TGAAGCCGGT GCTGCCAAGA
121 GCCATGCCCA GTTCCATGGG GGGTGGGGGT GGAGGCAGCC CCAGCCCTGT GGAGCTACGG
181 GGGGCTCTGG TGGGCTCTGT GGACCCCACA CTGCGGGAGC AGCAACTGCA GCAGGAGCTC
241 CTGGCGCTCA AGCAGCAGCA GCAGCTGCAG AAGCAGCTCC TGTTCGCTGA GTTCCAGAAA
301 CAGCATGACC ACCTGACAAG GCAGCATGAG GTCCAGCTGC AGAAGCACCT CAAGCAGCAG

361 CAGGAGATGC TGGCAGCCAA GCAGCAGCAG GAGATGCTGG CAGCCAAGCG GCAGCAGGAG
421 CTGGAGCAGC AGCGGCAGCG GGAGCAGCAG CGGCAGGAAG AGCTGGAGAA GCAGCGGCTG
481 GAGCAGCAGC TGCTCATCCT GCGGAACAAG GAGAAGAGCA AAGAGAGTGC CATTGCCAGC
541 ACTGAGGTAA AGCTGAGGCT CCAGGAATTC CTCTTGTCGA AGTCAAAGGA GCCCACACCA
601 GGCGGCCTCA ACCATTCCCT CCCACAGCAC CCCAAATGCT GGGGAGCCCA CCATGCTTCT
661 TTGGACCAGA GTTCCCCTCC CCAGAGCGGC CCCCCTGGGA CGCCTCCCTC CTACAAACTG
721 CCTTTGCCTG GGCCCTACGA CAGTCGAGAC GACTTCCCCC TCCGCAAAAC AGCCTCTGAA
781 CCCAACTTGA AAGTGCGTTC AAGGCTAAAA CAGAAGGTGG CTGAGCGGAG AAGCAGTCCC
841 CTCCTGCGTC GCAAGGATGG GACTGTTATT AGCACCTTTA AGAAGAGAGC TGTTGAGATC
901 ACAGGTGCCG GGCCTGGGGC GTCGTCCGTG TGTAACAGCG CACCCGGCTC CGGCCCCAGC
961 TCTCCCAACA GCTCCCACAG CACCATCGCT GAGAATGGCT TTACTGGCTC AGTCCCCAAC
1021 ATCCCCACTG AGATGCTCCC TCAGCACCGA GCCCTCCCTC TGGACAGCTC CCCCAACCAG
1081 TTCAGCCTCT ACACGTCTCC TTCTCTGCCC AACATCTCCC TAGGGCTGCA GGCCACGGTC
1141 ACTGTCACCA ACTCACACCT CACTGCCTCC CCGAAGCTGT CGACACAGCA GGAGGCCGAG
1201 AGGCAGGCCC TCCAGTCCCT GCGGCAGGGT GGCACGCTGA CCGGCAAGTT CATGAGCACA
1261 TCCTCTATTC CTGGCTGCCT GCTGGGCGTG GCACTGGAGG GCGACGGGAG CCCCCACGGG
1321 CATGCCTCCC TGCTGCAGCA TGTGCTGTTG CTGGAGCAGG CCCGGCAGCA GAGCACCCTC
1381 ATTGCTGTGC CACTCCACGG GCAGTCCCCA CTAGTGACGG GTGAACGTGT GGCCACCAGC
1441 ATGCGGACGG TAGGCAAGCT CCCGCGGCAT CGGCCCCTGA GCCGCACTCA GTCCTCACCG
1501 CTGCCGCAGA GTCCCCAGGC CCTGCAGCAG CTGGTCATGC AACAACAGCA CCAGCAGTTC
1561 CTGGAGAAGC AGAAGCAGCA GCAGCTACAG CTGGGCAAGA TCCTCACCAA GACAGGGGAG
1621 CTGCCCAGGC AGCCCACCAC CCACCCTGAG GAGACAGAGG AGGAGCTGAC GGAGCAGCAG
1681 GAGGTCTTGC TGGGGGAGGG AGCCCTGACC ATGCCCCGGG AGGGCTCCAC AGAGAGTGAG
1741 AGCACACAGG AAGACCTGGA GGAGGAGGAC GAGGAAGAGG ATGGGGAGGA GGAGGAGGAT
1801 TGCATCCAGG TTAAGGACGA GGAGGGCGAG AGTGGTGCTG AGGAGGGGCC CGACTTGGAG
1861 GAGCCTGGTG CTGGATACAA AAAACTGTTC TCAGATGCCC AGCCGCTGCA GCCTTTGCAG
1921 GTGTACCAGG CGCCCCTCAG CCTGGCCACT GTGCCCCACC AGGCCCTGGG CCGTACCCAG
1981 TCCTCCCCTG CTGCCCCTGG GGGCATGAAG AGCCCCCCAG ACCAGCCCGT CAAGCACCTC
2041 TTCACCACAG GTGTGGTCTA CGACACGTTC ATGCTAAAGC ACCAGTGCAT GTGCGGGAAC

2101 ACACACGTGC ACCCTGAGCA TGCTGGCCGG ATCCAGAGCA TCTGGTCCCG GCTGCAGGAG

2161 ACAGGCCTGC TTAGCAAGTG CGAGCGGATC CGAGGTCGCA AAGCCACGCT AGATGAGATC

2221 CAGACAGTGC ACTCTGAATA CCACACCCTG CTCTATGGGA CCAGTCCCCT CAACCGGCAG

2281 AAGCTAGACA GCAAGAAGTT GCTCGGCCCC ATCAGCCAGA AGATGTATGC TGTGCTGCCT

2341 TGTGGGGGCA TCGGGGTGGA CAGTGACACC GTGTGGAATG AGATGCACTC CTCCAGTGCT

2401 GTGCGCATGG CAGTGGGCTG CCTGCTGGAG CTGGCCTTCA AGGTGGCTGC AGGAGAGCTC

2461 AAGAATGGAT TTGCCATCAT CCGGCCCCCA GGACACCACG CCGAGGAATC CACAGCCATG

2521 GGATTCTGCT TCTTCAACTC TGTAGCCATC ACCGCAAAAC TCCTACAGCA GAAGTTGAAC

2581 GTGGGCAAGG TCCTCATCGT GGACTGGGAC ATTCACCATG GCAATGGCAC CCAGCAGGCG

2641 TTCTATAATG ACCCCTCTGT GCTCTACATC TCTCTGCATC GCTATGACAA CGGGAACTTC

2701 TTTCCAGGCT CTGGGGCTCC TGAAGAGGTT GGTGGAGGAC CAGGCGTGGG GTACAATGTG

2761 AACGTGGCAT GGACAGGAGG TGTGGACCCC CCCATTGGAG ACGTGGAGTA CCTTACAGCC

2821 TTCAGGACAG TGGTGATGCC CATTGCCCAC GAGTTCTCAC CTGATGTGGT CCTAGTCTCC

2881 GCCGGGTTTG ATGCTGTTGA AGGACATCTG TCTCCTCTGG GTGGCTACTC TGTCACCGCC

2941 AGATGTTTTG GCCACTTGAC CAGGCAGCTG ATGACCCTGG CAGGGGGCCG GGTGGTGCTG

3001 GCCCTGGAGG GAGGCCATGA CTTGACCGCC ATCTGTGATG CCTCTGAGGC TTGTGTCTCG

3061 GCTCTGCTCA GTGTAGAGCT GCAGCCCTTG GATGAGGCAG TCTTGCAGCA AAAGCCCAAC

3121 ATCAACGCAG TGGCCACGCT AGAGAAAGTC ATCGAGATCC AGAGCAAACA CTGGAGCTGT

3181 GTGCAGAAGT TCGCCGCTGG TCTGGGCCGG TCCCTGCGAG AGGCCCAAGC AGGTGAGACC

3241 GAGGAGGCCG AGACTGTGAG CGCCATGGCC TTGCTGTCGG TGGGGGCCGA GCAGGCCCAG

3301 GCTGCGGCAG CCCGGGAACA CAGCCCCAGG CCGGCAGAGG AGCCCATGGA GCAGGAGCCT

3361 GCCCTGTGA

SEQ. ID. NO. 4 (HDAC5)

MNSPNESDGMSGREPSLEILPRTSLHSIPVTVEVKPVLPRAMPSSMGGGGGGSPSP

VELRGALVGSVDPTLREQQLQQELLALKQQQQLQKQLLFAEFQKQHDHLTRQH

EVQLQKHLKQQQEMLAAKQQQEMLAAKRQQELEQQRQREQQRQEELEKQRLE

QQLLILRNKEKSKESAIASTEVKLRLQEFLLSKSKEPTPGGLNHSLPQHPKCWGA

HHASLDQSSPPQSGPPGTPPSYKLPLPGPYDSRDDFPLRKTASEPNLKVRSRLKQK

VAERRSSPLLRRKDGTVISTFKKRAVEITGAGPGASSVCNSAPGSGPSSPNSSHSTI

AENGFTGSVPNIPTEMLPQHRALPLDSSPNQFSLYTSPSLPNISLGLQATVTVTNS

HLTASPKLSTQQEAERQALQSLRQGGTLTGKFMSTSSIPGCLLGVALEGDGSPHG

HASLLQHVLLLEQARQQSTLIAVPLHGQSPLVTGERVATSMRTVGKLPRHRPLSR

TQSSPLPQSPQALQQLVMQQQHQQFLEKQKQQQLQLGKILTKTGELPRQPTTHP

EETEEELTEQQEVLLGEGALTMPREGSTESESTQEDLEEEDEEEDGEEEEDCIQVK

DEEGESGAEEGPDLEEPGAGYKKLFSDAQPLQPLQVYQAPLSLATVPHQALGRT

QSSPAAPGGMKSPPDQPVKHLFTTGVVYDTFMLKHQCMCGNTHVHPEHAGRIQ

SIWSRLQETGLLSKCERIRGRKATLDEIQTVHSEYHTLLYGTSPLNRQKLDSKKLL

GPISQKMYAVLPCGGIGVDSDTVWNEMHSSSAVRMAVGCLLELAFKVAAGELK

NGFAIIRPPGHHAEESTAMGFCFFNSVAITAKLLQQKLNVGKVLIVDWDIHHGNG

TQQAFYNDPSVLYISLHRYDNGNFFPGSGAPEEVGGGPGVGYNVNVAWTGGVD

PPIGDVEYLTAFRTVVMPIAHEFSPDVVLVSAGFDAVEGHLSPLGGYSVTARCFG

HLTRQLMTLAGGRVVLALEGGHDLTAICDASEACVSALLSVELQPLDEAVLQQK

PNINAVATLEKVIEIQSKHWSCVQKFAAGLGRSLREAQAGETEEAETVSAMALLS

VGAEQAQAAAAREHSPRPAEEPMEQEPAL

SEQ. ID. NO. 5 (HDAC6)

1 ATGACCTCAA CCGGCCAGGA TTCCACCACA ACCAGGCAGC GAAGAAGTAG GCAGAACCCC

61 CAGTCGCCCC CTCAGGACTC CAGTGTCACT TCGAAGCGAA ATATTAAAAA GGGAGCCGTT

121 CCCCGCTCTA TCCCCAATCT AGCGGAGGTA AAGAAGAAAG GCAAAATGAA GAAGCTCGGC

181 CAAGCAATGG AAGAAGACCT AATCGTGGGA CTGCAAGGGA TGGATCTGAA CCTTGAGGCT

241 GAAGCACTGG CTGGCACTGG CTTGGTGTTG GATGAGCAGT TAAATGAATT CCATTGCCTC

301 TGGGATGACA GCTTCCCGGA AGGCCCTGAG CGGCTCCATG CCATCAAGGA GCAACTGATC
361 CAGGAGGGCC TCCTAGATCG CTGCGTGTCC TTTCAGGCCC GGTTTGCTGA AAAGGAAGAG
421 CTGATGTTGG TTCACAGCCT AGAATATATT GATCTGATGG AAACAACCCA GTACATGAAT
481 GAGGGAGAAC TCCGTGTCCT AGCAGACACC TACGACTCAG TTTATCTGCA TCCGAACTCA
541 TACTCCTGTG CCTGCCTGGC CTCAGGCTCT GTCCTCAGGC TGGTGGATGC GGTCCTGGGG
601 GCTGAGATCC GGAATGGCAT GGCCATCATT AGGCCTCCTG GACATCACGC CCAGCACAGT
661 CTTATGGATG GCTATTGCAT GTTCAACCAC GTGGCTGTGG CAGCCCGCTA TGCTCAACAG
721 AAACACCGCA TCCGGAGGGT CCTTATCGTA GATTGGGATG TGCACCACGG TCAAGGAACA
781 CAGTTCACCT TCGACCAGGA CCCCAGTGTC CTCTATTTCT CCATCCACCG CTACGAGCAG
841 GGTAGGTTCT GGCCCCACCT GAAGGCCTCT AACTGGTCCA CCACAGGTTT CGGCCAAGGC
901 CAAGGATATA CCATCAATGT GCCTTGGAAC CAGGTGGGGA TGCGGGATGC TGACTACATT
961 GCTGCTTTCC TGCACGTCCT GCTGCCAGTC GCCCTCGAGT TCCAGCCTCA GCTGGTCCTG
1021 GTGGCTGCTG GATTTGATGC CCTGCAAGGG GACCCCAAGG GTGAGATGGC CGCCACTCCG
1081 GCAGGGTTCG CCCAGCTAAC CCACCTGCTC ATGGGTCTGG CAGGAGGCAA GCTGATCCTG
1141 TCTCTGGAGG GTGGCTACAA CCTCCGCGCC CTGGCTGAAG GCGTCAGTGC TTCGCTCCAC
1201 ACCCTTCTGG GAGACCCTTG CCCCATGCTG GAGTCACCTG GTGCCCCCTG CCGGAGTGCC
1261 CAGGCTTCAG TTTCCTGTGC TCTGGAAGCC CTTGAGCCCT TCTGGGAGGT TCTTGTGAGA
1321 TCAACTGAGA CCGTGGAGAG GGACAACATG GAGGAGGACA ATGTAGAGGA GAGCGAGGAG
1381 GAAGGACCCT GGGAGCCCCC TGTGCTCCCA ATCCTGACAT GGCCAGTGCT ACAGTCTCGC
1441 ACAGGGCTGG TCTATGACCA AAATATGATG AATCACTGCA ACTTGTGGGA CAGCCACCAC
1501 CCTGAGGTAC CCCAGCGCAT CTTGCGGATC ATGTGCCGTC TGGAGGAGCT GGGCCTTGCC
1561 GGGCGCTGCC TCACCCTGAC ACCGCGCCCT GCCACAGAGG CTGAGCTGCT CACCTGTCAC
1621 AGTGCTGAGT ACGTGGGTCA TCTCCGGGCC ACAGAGAAAA TGAAAACCCG GGAGCTGCAC
1681 CGTGAGAGTT CCAACTTTGA CTCCATCTAT ATCTGCCCCA GTACCTTCGC CTGTGCACAG
1741 CTTGCCACTG GCGCTGCCTG CCGCCTGGTG GAGGCTGTGC TCTCAGGAGA GGTTCTGAAT
1801 GGTGCTGCTG TGGTGCGTCC CCCAGGACAC CACGCAGAGC AGGATGCAGC TTGCGGTTTT
1861 TGCTTTTTCA ACTCTGTGGC TGTGGCTGCT CGCCATGCCC AGACTATCAG TGGGCATGCC
1921 CTACGGATCC TGATTGTGGA TTGGGATGTC CACCACGGTA ATGGAACTCA GCACATGTTT
1981 GAGGATGACC CCAGTGTGCT ATATGTGTCC CTGCACCGCT ATGATCATGG CACCTTCTTC

2041 CCCATGGGGG ATGAGGGTGC CAGCAGCCAG ATCGGCCGGG CTGCGGGCAC AGGCTTCACC

2101 GTCAACGTGG CATGGAACGG GCCCCGCATG GGTGATGCTG ACTACCTAGC TGCCTGGCAT

2161 CGCCTGGTGC TTCCCATTGC CTACGAGTTT AACCCAGAAC TGGTGCTGGT CTCAGCTGGC

2221 TTTGATGCTG CACGGGGGGA TCCGCTGGGG GGCTGCCAGG TGTCACCTGA GGGTTATGCC

2281 CACCTCACCC ACCTGCTGAT GGGCCTTGCC AGTGGCCGCA TTATCCTTAT CCTAGAGGGT

2341 GGCTATAACC TGACATCCAT CTCAGAGTCC ATGGCTGCCT GCACTCGCTC CCTCCTTGGA

2401 GACCCACCAC CCCTGCTGAC CCTGCCACGG CCCCCACTAT CAGGGGCCCT GGCCTCAATC

2461 ACTGAGACCA TCCAAGTCCA TCGCAGATAC TGGCGCAGCT TACGGGTCAT GAAGGTAGAA

2521 GACAGAGAAG GACCCTCCAG TTCTAAGTTG GTCACCAAGA AGGCACCCCA ACCAGCCAAA

2581 CCTAGGTTAG CTGAGCGGAT GACCACACGA GAAAAGAAGG TTCTGGAAGC AGGCATGGGG

2641 AAAGTCACCT CGGCATCATT TGGGGAAGAG TCCACTCCAG GCCAGACTAA CTCAGAGACA

2701 GCTGTGGTGG CCCTCACTCA GGACCAGCCC TCAGAGGCAG CCACAGGGGG AGCCACTCTG

2761 GCCCAGACCA TTTCTGAGGC AGCCATTGGG GGAGCCATGC TGGGCCAGAC CACCTCAGAG

2821 GAGGCTGTCG GGGGAGCCAC TCCGGACCAG ACCACCTCAG AGGAGACTGT GGGAGGAGCC

2881 ATTCTGGACC AGACCACCTC AGAGGATGCT GTTGGGGGAG CCACGCTGGG CCAGACTACC

2941 TCAGAGGAGG CTGTAGGAGG AGCTACACTG GCCCAGACCA TCTCGGAGGC AGCCATGGAG

3001 GGAGCCACAC TGGACCAGAC TACGTCAGAG GAGGCTCCAG GGGGCACCGA GCTGATCCAA

3061 ACTCCTCTAG CCTCGAGCAC AGACCACCAG ACCCCCCCAA CCTCACCTGT GCAGGGAACT

3121 ACACCCCAGA TATCTCCCAG TACACTGATT GGGAGTCTCA GGACCTTGGA GCTAGGCAGC

3181 GAATCTCAGG GGGCCTCAGA ATCTCAGGCC CCAGGAGAGG AGAACCTACT AGGAGAGGCA

3241 GCTGGAGGTC AGGACATGGC TGATTCGATG CTGATGCAGG GATCTAGGGG CCTCACTGAT

3301 CAGGCCATAT TTTATGCTGT GACACCACTG CCCTGGTGTC CCCATTTGGT GGCAGTATGC

3361 CCCATACCTG CAGCAGGCCT AGACGTGACC CAACCTTGTG GGGACTGTGG AACAATCCAA

3421 GAGAATTGGG TGTGTCTCTC TTGCTATCAG GTCTACTGTG GTCGTTACAT CAATGGCCAC

3481 ATGCTCCAAC ACCATGGAAA TTCTGGACAC CCGCTGGTCC TCAGCTACAT CGACCTGTCA

3541 GCCTGGTGTT ACTACTGTCA GGCCTATGTC CACCACCAGG CTCTCCTAGA TGTGAAGAAC

3601 ATCGCCCACC AGAACAAGTT TGGGGAGGAT ATGCCCCACC CACACTAA

SEQ ID. NO. 6 (HDAC6)

MTSTGQDSTTTRQRRSRQNPQSPPQDSSVTSKRNIKKGAVPRSIPNLAEVKKKGK
MKKLGQAMEEDLIVGLQGMDLNLEAEALAGTGLVLDEQLNEFHCLWDDSFPEG
PERLHAIKEQLIQEGLLDRCVSFQARFAEKEELMLVHSLEYIDLMETTQYMN
EGELRVLADTYDSVYLHPNSYSCACLASGSVLRLVDAVLGAEIRNGMAIIRPPGH
HAQHSLMDGYCMFNHVAVAARYAQQKHRIRRVLIVDWDVHHGQGTQFTFDQD
PSVLYFSIHRYEQGRFWPHLKASNWSTTGFGQGQGYTINVPWNQVGMRDADYI
AAFLHVLLPVALEFQPQLVLVAAGFDALQGDPKGEMAATPAGFAQLTHLLMGL
AGGKLILSLEGGYNLRALAEGVSASLHTLLGDPCPMLESPGAPCRSAQASVSCAL
EALEPFWEVLVRSTETVERDNMEEDNVEESEEEGPWEPPVLPILTWPVLQSRTGL
VYDQNMMNHCNLWDSHHPEVPQRILRIMCRLEELGLAGRCLTLTPRPATEAELL
TCHSAEYVGHLRATEKMKTRELHRESSNFDSIYICPSTFACAQLATGAACRLVEA
VLSGEVLNGAAVVRPPGHHAEQDAACGFCFFNSVAVAARHAQTISGHALRILIV
DWDVHHGNGTQHMFEDDPSVLYVSLHRYDHGTFFPMGDEGASSQIGRAAGTGF
TVNVAWNGPRMGDADYLAAWHRLVLPIAYEFNPELVLVSAGFDAARGDPLGG
CQVSPEGYAHLTHLLMGLASGRIILILEGGYNLTSISESMAACTRSLLGDPPPLLTL
PRPPLSGALASITETIQVHRRYWRSLRVMKVEDREGPSSSKLVTKKAPQPAKPRL
AERMTTREKKVLEAGMGKVTSASFGEESTPGQTNSETAVVALTQDQPSEAA
TGGATLAQTISEAAIGGAMLGQTTSEEAVGGATPDQTTSEETVGGAILDQTTSED
AVGGATLGQTTSEEAVGGATLAQTISEAAMEGATLDQTTSEEAPGGTELIQTPLA
SSTDHQTPPTSPVQGTTPQISPSTLIGSLRTLELGSESQGASESQAPGEENLLGEAA
GGQDMADSMLMQGSRGLTDQAIFYAVTPLPWCPHLVAVCPIPAAGLDVTQPCG

DCGTIQENWVCLSCYQVYCGRYINGHMLQHHGNSGHPLVLSYIDLSAWCYYCQ

AYVHHQALLDVKNIAHQ NKFGEDMPHPH

SEQ. ID. NO. 7

HHAXXXXXXGXCXFNXVAXXAXXXQXXXXXXXXXXXLIVDWDXHHGXGTQX

XFXXDPSVLYXSXHRYXXGXFXPXX

SEQ. ID. NO. 8

HHAEESTPMGFCYFNSVAVAAKLLQQRLSVSKILIVDWDVHHGNGTQQAFYSDP

SVLYMSLHRYDDGNFFPGS

SEQ. ID. NO. 9

HHAEESTAMGFCFFNSVAITAKLLQQKLNVGKVLIVDWDIHHGNGTQQAFYNDP

SVLYISLHRYDNGNFFPGS

SEQ. ID. NO. 10

HHAQHSLMDGYCMFNHVAVAARYAQQKHRIRRVLIVDWDVHHGQGTQFTFDQ

DPSVLYFSIHRYEQGRFWPHL

SEQ. ID. NO. 11 (HDAC7A)

```
1    ATAATACCTA CCTTGCAGGA CCACGACAGG ATTAAGTGAG GAAAAACCCC CATGAGAGTG
61   TTTTGCCATT GTCAAGTGAG CCTGAGGGAG GCTGAGGGGG GATCAGGCTG TATCATGCCC
121  CCGAGGACAA ACTTTCCAGT TTACCCTGCT CCCTCTCTCT GTCCCTAGGC TGCCCCAGGC
181  CCTGTGCAGA CACACCAGGC CCTCAGCCGC AGCCCATGGA CCTGCGGGTG GGCCAGCGGC
```

241 CCCCAGTGGA GCCCCCACCA GAGCCCACAT TGCTGGCCCT GCAGCGTCCC CAGCGCCTGC

301 ACCACCACCT CTTCCTAGCA GGCCTGCAGC AGCAGCGCTC GGTGGAGCCC ATGAGGCTCT

361 CCATGGACAC GCCGATGCCC GAGTTGCAGG TGGGACCCCA GGAACAAGAG CTGCGGCAGC

421 TTCTCCACAA GGACAAGAGC AAGCGAAGTG CTGTAGCCAG CAGCGTGGTC AAGCAGAAGC

481 TAGCGGAGGT GATTCTGAAA AAACAGCAGG CGGCCCTAGA AAGAACAGTC CATCCCAACA

541 GCCCCGGCAT TCCCTACAGA ACCCTGGAGC CCCTGGAGAC GGAAGGAGCC ACCCGCTCCA

601 TGCTCAGCAG CTTTTTGCCT CCTGTTCCCA GCCTGCCCAG TGACCCCCCA GAGCACTTCC

661 CTCTGCGCAA GACAGTCTCT GAGCCCAACC TGAAGCTGCG CTATAAGCCC AAGAAGTCCC

721 TGGAGCGGAG GAAGAATCCA CTGCTCCGAA AGGAGAGTGC GCCCCCCAGC CTCCGGCGGC

781 GGCCCGCAGA GACCCTCGGA GACTCCTCCC CAAGTAGTAG CAGCACGCCC GCATCAGGGT

841 GCAGCTCCCC CAATGACAGC GAGCACGGCC CCAATCCCAT CCTGGGCGAC AGTGACCGCA

901 GGACCCATCC GACTCTGGGC CCTCGGGGGC CAATCCTGGG GAGCCCCCAC ACTCCCCTCT

961 TCCTGCCCCA TGGCTTGGAG CCCGAGGCTG GGGGCACCTT GCCCTCTCGC CTGCAGCCCA

1021 TTCTCCTCCT GGACCCCTCA GGCTCTCATG CCCCGCTGCT GACTGTGCCC GGGCTTGGGC

1081 CCTTGCCCTT CCACTTTGCC CAGTCCTTAA TGACCACCGA GCGGCTCTCT GGGTCAGGCC

1141 TCCACTGGCC ACTGAGCCGG ACTCGCTCAG AGCCCCTGCC CCCCAGTGCC ACCGCTCCCC

1201 CACCGCCGGG CCCCATGCAG CCCCGCCTGG AGCAGCTCAA AACTCACGTC CAGGTGATCA

1261 AGAGGTCAGC CAAGCCGAGT GAGAAGCCCC GGCTGCGGCA GATACCCTCG GCTGAAGACC

1321 TGGAGACAGA TGGCGGGGGA CCGGGCCAGG TGGTGGACGA TGGCCTGGAG CACAGGGAGC

1381 TGGGCCATGG GCAGCCTGAG GCCAGAGGCC CCGCTCCTCT CCAGCAGCAC CCTCAGGTGT

1441 TGCTCTGGGA ACAGCAGCGA CTGGCTGGGC GGCTCCCCCG GGGCAGCACC GGGGACACTG

1501 TGCTGCTTCC TCTGGCCCAG GGTGGGCACC GGCCTCTGTC CCGGGCTCAG TCTTCCCCAG

1561 CCGCACCTGC CTCACTGTCA GCCCCAGAGC CTGCCAGCCA GGCCCGAGTC CTCTCCAGCT

1621 CAGAGACCCC TGCCAGGACC CTGCCCTTCA CCACAGGGCT GATCTATGAC TCGGTCATGC

1681 TGAAGCACCA GTGCTCCTGC GGTGACAACA GCAGGCACCC GGAGCACGCC GGCCGCATCC

1741 AGAGCATCTG GTCCCGGCTG CAGGAGCGGG GGCTCCGGAG CCAGTGTGAG TGTCTCCGAG

1801 GCCGGAAGGC CTCCCTGGAA GAGCTGCAGT CGGTCCACTC TGAGCGGCAC GTGCTCCTCT

1861 ACGGCACCAA CCCGCTCAGC CGCCTCAAAC TGGACAACGG GAAGCTGGCA GGGCTCCTGG

1921 CACAGCGGAT GTTTGTGATG CTGCCCTGTG GTGGGGTTGG GGTGGACACT GACACCATCT

1981 GGAATGAGCT TCATTCCTCC AATGCAGCCC GCTGGGCCGC TGGCAGTGTC ACTGACCTCG

2041 CCTTCAAAGT GGCTTCTCGT GAGCTAAAGA ATGGTTTCGC TGTGGTGCGG CCCCCAGGAC

2101 ACCATGCAGA TCATTCAACA GCCATGGGCT TCTGCTTCTT CAACTCAGTG GCCATCGCCT

2161 GCCGGCAGCT GCAACAGCAG AGCAAGGCCA GCAAGATCCT CATTGTAGAC TGGGACGTGC

2221 ACCATGGCAA CGGCACCCAG CAAACCTTCT ACCAAGACCC CAGTGTGCTC TACATCTCCC

2281 TGCATCGCCA TGACGACGGC AACTTCTTCC CGGGGAGTGG GGCTGTGGAT GAGGTAGGGG

2341 CTGGCAGCGG TGAGGGCTTC AATGTCAATG TGGCCTGGGC TGGAGGTCTG GACCCCCCCA

2401 TGGGGGATCC TGAGTACCTG GCTGCTTTCA GGATAGTCGT GATGCCCATC GCCCGAGAGT

2461 TCTCTCCAGA CCTAGTCCTG GTGTCTGCTG GATTTGATGC TGCTGAGGGT CACCCGGCCC

2521 CACTGGGTGG CTACCATGTT TCTGCCAAAT GTTTTGGATA CATGACGCAG CAACTGATGA

2581 ACCTGGCAGG AGGCGCAGTG GTGCTGGCCT TGGAGGGTGG CCATGACCTC ACAGCCATCT

2641 GTGACGCCTC TGAGGCCTGT GTGGCTGCTC TTCTGGGTAA CAGGGTGGAT CCCCTTTCAG

2701 AAGAAGGCTG GAAACAGAAA CCCAACCTCA ATGCCATCCG CTCTCTGGAG GCCGTGATCC

2761 GGGTGCACAG TAAATACTGG GGCTGCATGC AGCGCCTGGC CTCCTGTCCA GACTCCTGGG

2821 TGCCTAGAGT GCCAGGGGCT GACAAAGAAG AAGTGGAGGC AGTGACCGCA CTGGCGTCCC

2881 TCTCTGTGGG CATCCTGGCT GAAGATAGGC CCTCGGAGCA GCTGGTGGAG GAGGAAGAAC

2941 CTATGAATCT CTAAGGCTCT GGAACCATCT GCCCGCCCAC CATGCCCTTG GGACCTGGTT

3001 CTCTTCTAAC CCCTGGCAAT AGCCCCCATT CCTGGGTCTT TAGAGATCCT GTGGGCAAGT

3061 AGTTGGAACC AGAGAACAGC CTGCCTGCTT TGACAGTTAT CCCAGGGAGC GTGAGAAAAT

3121 C

SEQ. ID. NO. 12 (HDAC7A)

1 MDLRVGQRPP VEPPPEPTLL ALQRPQRLHH HLFLAGLQQQ RSVEPMRLSM DTPMPELQVG

61 PQEQELRQLL HKDKSKRSAV ASSVVKQKLA EVILKKQQAA LERTVHPNSP GIPYRTLEPL

121 ETEGATRSML SSFLPPVPSL PSDPPEHFPL RKTVSEPNLK LRYKPKKSLE RRKNPLLRKE

181 SAPPSLRRRP AETLGDSSPS SSSTPASGCS SPNDSEHGPN PILGDSDRRT HPTLGPRGPI

241 LGSPHTPLFL PHGLEPEAGG TLPSRLQPIL LLDPSGSHAP LLTVPGLGPL PFHFAQSLMT

301 TERLSGSGLH WPLSRTRSEP LPPSATAPPP PGPMQPRLEQ LKTHVQVIKR SAKPSEKPRL

361 RQIPSAEDLE TDGGGPGQVV DDGLEHRELG HGQPEARGPA PLQQHPQVLL WEQQRLAGRL

421 PRGSTGDTVL LPLAQGGHRP LSRAQSSPAA PASLSAPEPA SQARVLSSSE TPARTLPFTT

481 GLIYDSVMLK HQCSCGDNSR HPEHAGRIQS IWSRLQERGL RSQCECLRGR KASLEELQSV

541 HSERHVLLYG TNPLSRLKLD NGKLAGLLAQ RMFVMLPCGG VGVDTDTIWN ELHSSNAARW

601 AAGSVTDLAF KVASRELKNG FAVVRPPGHH ADHSTAMGFC FFNSVAIACR QLQQQSKASK

661 ILIVDWDVHH GNGTQQTFYQ DPSVLYISLH RHDDGNFFPG SGAVDEVGAG SGEGFNVNVA

721 WAGGLDPPMG DPEYLAAFRI VVMPIAREFS PDLVLVSAGF DAAEGHPAPL GGYHVSAKCF

781 GYMTQQLMNL AGGAVVLALE GGHDLTAICD ASEACVAALL GNRVDPLSEE GWKQKPNLNA

841 IRSLEAVIRV HSKYWGCMQR LASCPDSWVP RVPGADKEEV EAVTALASLS VGILAEDRPS

901 EQLVEEEEPM NL

SEQ. ID. NO. 13 (HDAC7B)

1 GGGGAAGAGA GGCACAGACA CAGATAGGAG AAGGGCACCG GCTGGAGCCA CTTGCAGGAC

61 TGAGGGTTTT TGCAACAAAA CCCTAGCAGC CTGAAGAACT CTAAGCCAGA TGGGGTGGCT

121 GGACGAGAGC AGCTCTTGGC TCAGCAAAGA ATGCACAGTA TGATCAGCTC AGTGGATGTG

181 AAGTCAGAAG TTCCTGTGGG CCTGGAGCCC ATCTCACCTT TAGACCTAAG GACAGACCTC

241 AGGATGATGA TGCCCGTGGT GGACCCTGTT GTCCGTGAGA AGCAATTGCA GCAGGAATTA

301 CTTCTTATCC AGCAGCAGCA ACAAATCCAG AAGCAGCTTC TGATAGCAGA GTTTCAGAAA

361 CAGCATGAGA ACTTGACACG GCAGCACCAG GCTCAGCTTC AGGAGCATAT CAAGGAACTT

421 CTAGCCATAA AACAGCAACA AGAACTCCTA GAAAAGGAGC AGAAACTGGA GCAGCAGAGG

481 CAAGAACAGG AAGTAGAGAG GCATCGCAGA GAACAGCAGC TTCCTCCTCT CAGAGGCAAA

541 GATAGAGGAC GAGAAAGGGC AGTGGCAAGT ACAGAAGTAA AGCAGAAGCT TCAAGAGTTC

601 CTACTGAGTA AATCAGCAAC GAAAGACACT CCAACTAATG GAAAAAATCA TTCCGTGAGC

661 CGCCATCCCA AGCTCTGGTA CACGGCTGCC CACCACACAT CATTGGATCA AAGCTCTCCA

721 CCCCTTAGTG GAACATCTCC ATCCTACAAG TACACATTAC CAGGAGCACA AGATGCAAAG

781 GATGATTTCC CCCTTCGAAA AACTGCCTCT GAGCCCAACT TGAAGGTGCG GTCCAGGTTA

841 AAACAGAAAG TGGCAGAGAG GAGAAGCAGC CCCTTACTCA GGCGGAAGGA TGGAAATGTT

901 GTCACTTCAT TCAAGAAGCG AATGTTTGAG GTGACAGAAT CCTCAGTCAG TAGCAGTTCT

961 CCAGGCTCTG GTCCCAGTTC ACCAAACAAT GGGCCAACTG GAAGTGTTAC TGAAAATGAG

1021 ACTTCGGTTT TGCCCCCTAC CCCTCATGCC GAGCAAATGG TTTCACAGCA ACGCATTCTA

1081 ATTCATGAAG ATTCCATGAA CCTGCTAAGT CTTTATACCT CTCCTTCTTT GCCCAACATT

1141 ACCTTGGGGC TTCCCGCAGT GCCATCCCAG CTCAATGCTT CGAATTCACT CAAAGAAAAG

1201 CAGAAGTGTG AGACGCAGAC GCTTAGGCAA GGTGTTCCTC TGCCTGGGCA GTATGGAGGC

1321 AGCCACCAGG CTCTCCTGCA GCATTTATTA TTGAAAGAAC AAATGCGACA GCAAAAGCTT

1381 CTTGTAGCTG GTGGAGTTCC CTTACATCCT CAGTCTCCCT TGGCAACAAA AGAGAGAATT

1441 TCACCTGGCA TTAGAGGTAC CCACAAATTG CCCCGTCACA GACCCCTGAA CCGAACCCAG

1501 TCTGCACCTT TGCCTCAGAG CACGTTGGCT CAGCTGGTCA TTCAACAGCA ACACCAGCAA

1561 TTCTTGGAGA AGCAGAAGCA ATACCAGCAG CAGATCCACA TGAACAAACT GCTTTCGAAA

1621 TCTATTGAAC AACTGAAGCA ACCAGGCAGT CACCTTGAGG AAGCAGAGGA AGAGCTTCAG

1681 GGGGACCAGG CGATGCAGGA AGACAGAGCG CCCTCTAGTG GCAACAGCAC TAGGAGCGAC

1741 AGCAGTGCTT GTGTGGATGA CACACTGGGA CAAGTTGGGG CTGTGAAGGT CAAGGAGGAA

1801 CCAGTGGACA GTGATGAAGA TGCTCAGATC CAGGAAATGG AATCTGGGGA GCAGGCTGCT

1861 TTTATGCAAC AGGTAATAGG CAAAGATTTA GCTCCAGGAT TTGTAATTAA AGTCATTATC

1921 TGAACATGAA ATGCATTGCA GGTTTGGTAA ATGGATATGA TTTCCTATCA GTTTATATTT

1981 CTCTATGATT TGAGTTCAGT GTTTAAGGAT TCTACCTAAT GCAGATATAT GTATATATCT

2041 ATATAGAGGT CTTTCTATAT ACTGATCTCT ATATAGATAT CAATGTTTCA TTGAAAATCC

2101 ACTGGTAAGG AAATACCTGT TATACTAAAA TTATGATACA TAATATCTGA GCAGTTAATA

2161 GGCTTTAAAT TTATCCCAAA GCCTGCTACA CCAATTACTT CTAAAGAAAA CAAATTCACT

2221 GTTATTTTGA GTTTATGTGT TGAGATCAGT GACTGCTGGA TAGTCTCCCA GTCTGATCAA

2281 TGAAGCATTC GATTAGTTTT TGATTTTTTG CAACATCTAG AATTTAATTT TCACATCACT

2341 GTACATAATG TATCATACTA TAGTCTTGAA CACTGTTAAA GGTAGTCTGC CCCTTCCTTC

2401 CTCTCTCTTT TTTTAGTTAA GTAGAAATGT TCTGGTCACC ATGCCAGTAG TCCTAGGTTA

2461 TTGTGTAGGT TGCAATTGAA CATATTAGGA ATACAGGTGG TTTTAAATAT ATAGATGCAA

2521 ATTGCAGCAC TACTTTAAAT ATTAGATTAT GTCTCACATA GCACTGCTCA TTTTACTTTT

2581 ATTTTGTGTA ATTTGATGAC ACTGTCTATC AAAAAAGAGC AAATGAAGCA GATGCAAATG

2641 TTAGTGAGAA GTAATGTGCA GCATTATGGT CCAATCAGAT ACAATATTGT GTCTACAATT

2701 GCAAAAAACA CAGTAACAGG ATGAATATTA TCTGATATCA AGTCAAAATC AGTTTGAAAA

2761 GAAGGTGTAT CATATTTTAT ATTGTCACTA GAATCTCTTA AGTATAATTC CATAATGACA

2821 TGGGCATATA CCGTAACATT CTGGCAAATA ACAATTAGAA AAGATAGGTT TAACAAAAAA

2881 ATTTACTTGT ATATAATGCA CCTTCAGGAG GACTATGTCC TTTGATGCTA TAAAATACAA

2941 ACAACTTTGA AGGCAACAGA AGACACTGTT TATTCAAGTC AGTTCTTTGT CAGGTTCCTG

3001 CTGTTCTCCT ACAGAAAAGT GATTCTGTGA GGGTGAACAG GAAATGCCTT GTGGAAACAG

3061 GAAGTCCAAG TGATTCATGT ACTGAGGAAT GTAGGAAAAA AAATCTGAGG ATAGTGCTTT

3121 ACTCTTTCTG TTTTTAAAGG GCACTCTATG AATTGATTTA TTGTCTAAGA AAATAACACC

3181 ACAAGTAGGG AAATTGTTAC GGAAGCTTTT CACTGGAACA TTTCCTTCAT ATTCCCTTTT

3241 GATATGTTTA CCTTGTTTTA TAGGTTTACT TTTGTTAAGC TAGTTAAAGG TTCGTTGTAT

3301 TAAGACCCCT TTAATATGGA TAATCCAAAT TGACCTAGAA TCTTTGTGAG GTTTTTTCTA

3361 TTAAAATATT TATATTTCTA AATCCGAGGT ATTTCAAGGT GTAGTATCCT ATTTCAAAGG

3421 AGATATAGCA GTTTTGCCAA ATGTAGACAT TGTTCAACTG TATGTTATTG GCACGTGTTG

3481 TTTACATTTT GCTGTGACAT TTAAAAATAT TTCTTTAAAA ATGTTACTGC TAAAGATACA

3541 TTATCCTTTT TTAAAAAGTC TCCATTCAAA TTAAATTAAC ATAACTAGAA GTTAGAAAGT

3601 TTAAAAGTTT TCCACATAAT GAAAGTCCTT CTGATAATTT GACAAATAGC TATAATAGGA

3661 ACACTCCCTA TCACCAACAT ATTTTGGTTA GTATATTCCT TCATATTAAA ATGACTTTTT

3721 GTCAGTTGTT TTGCATTAAA AATATGGCAT GCCTAAGATA AAATTGTATA TTTTTTCCAT

3781 CTCATAAATA TTCATTTTCT TCAAAGTCTT TTTTCAATCT CATAAAAAAG GGATAGTGCA

3841 TCTTTTAAAA TACATTTTAT TTGGGGAGGA ACATGTGGCT GAGCAGACTT TTGTATAATA

3901 TTACTTCAAA GATATGTAAT CACAAACAAA AAAAACTATT TTTTATAATG TCATTTGAGA

3961 GAGTTTCATC AGTACAGTTG GTGGACGTTA ATTGTTTGAA TTTGATAGTC TTTGAATTTA

4021 ATCAAGAAAC TACCTGGAAC CAGTGAAAAG GAAAGCTGGA CTTAAATAAT CTTAGAATTA

4081 ATTGATAAAT GTCTCTTTTA AAATCTACTG TATTTATTAT AATTTACACC CTTGAAGGTG

4141 ATCTCTTGTT TTGTGTTGTA AATATATTGT TTGTATGTTT CCCTTCTTGC CTTCTGTTAT

4201 AAGTCTCTTC CTTTCTCAAA TAAAGTTTTT TTTAAAAG

SEQ. ID. NO. 14 (HDAC7B)

1 MHSMISSVDV KSEVPVGLEP ISPLDLRTDL RMMMPVVDPV VREKQLQQEL LLIQQQQQIQ

61 KQLLIAEFQK QHENLTRQHQ AQLQEHIKEL LAIKQQQELL EKEQKLEQQR QEQEVERHRR

121 EQQLPPLRGK DRGRERAVAS TEVKQKLQEF LLSKSATKDT PTNGKNHSVS RHPKLWYTAA

181 HHTSLDQSSP PLSGTSPSYK YTLPGAQDAK DDFPLRKTAS EPNLKVRSRL KQKVAERRSS

241 PLLRRKDGNV VTSFKKRMFE VTESSVSSSS PGSGPSSPNN GPTGSVTENE TSVLPPTPHA

301 EQMVSQQRIL IHEDSMNLLS LYTSPSLPNI TLGLPAVPSQ LNASNSLKEK QKCETQTLRQ

361 GVPLPGQYGG SIPASSSHPH VTLEGKPPNS SHQALLQHLL LKEQMRQQKL LVAGGVPLHP

421 QSPLATKERI SPGIRGTHKL PRHRPLNRTQ SAPLPQSTLA QLVIQQQHQQ FLEKQKQYQQ

481 QIHMNKLLSK SIEQLKQPGS HLEEAEEELQ GDQAMQEDRA PSSGNSTRSD SSACVDDTLG

541 QVGAVKVKEE PVDSDEDAQI QEMESGEQAA FMQQVIGKDL APGFVIKVII